

# Buffer-Aided Relaying with Adaptive Link Selection — Fixed and Mixed Rate Transmission

Nikola Zlatanov and Robert Schober

## Abstract

We consider a simple network consisting of a source, a half-duplex decode-and-forward relay with a buffer, and a destination. We assume that the direct source-destination link is not available and all links undergo fading. We propose two new buffer-aided relaying schemes with different requirements regarding the availability of channel state information at the transmitter (CSIT). In the first scheme, neither the source nor the relay have CSIT, and consequently, both nodes are forced to transmit with fixed rates. In contrast, in the second scheme, the source does not have CSIT and transmits with fixed rate but the relay has CSIT and adapts its transmission rate accordingly. In the absence of delay constraints, for both fixed rate and mixed rate transmission, we derive the throughput-optimal buffer-aided relaying protocols which select either the source or the relay for transmission based on the instantaneous signal-to-noise ratios (SNRs) of the source-relay and the relay-destination links. In addition, for the delay constrained case, we develop buffer-aided relaying protocols that achieve a predefined average delay. Compared to conventional relaying protocols, which select the transmitting node according to a predefined schedule independent of the link instantaneous SNRs, the proposed buffer-aided protocols with adaptive link selection achieve large performance gains. In particular, for fixed rate transmission, we show that the proposed protocol achieves a diversity gain of two as long as an average delay of more than three time slots can be afforded. Furthermore, for mixed rate transmission with an average delay of  $E\{T\}$  time slots, a multiplexing gain of  $r = 1 - 1/(2E\{T\})$  is achieved. As a by-product of the considered link adaptive protocols, we also develop a novel conventional relaying protocol for mixed rate transmission, which yields the same multiplexing gain as the protocol with adaptive link selection. Hence, for mixed rate transmission, for sufficiently large average delays, buffer-aided half-duplex relaying with and without adaptive link selection does not suffer from a multiplexing gain loss compared to full-duplex relaying.

This paper was presented in part at IEEE Globecom 2011.

N. Zlatanov and R. Schober are with the Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, BC, V6T 1Z4, Canada, E-mail: zlatanov@ece.ubc.ca, rschober@ece.ubc.ca

## I. INTRODUCTION

Node cooperation can introduce significant throughput and diversity gains in wireless networks. The relay channel was first investigated by van der Meulen [1]. Later Cover and El Gamal [2] investigated the memoryless three-node relay channel consisting of a source, a destination, and a single full-duplex relay and proved that cooperative systems offer throughput gains compared to non-cooperative systems. This work was later extended to systems employing a half-duplex relay in fading environments for the case when the relay has a predetermined schedule for reception and transmission [3]. For the case of fixed rate transmission, the outage probability of the three-node relay network was shown to be superior to non-relay aided transmission in [4], [5]. Subsequently, in [6], a simple protocol for the three-node relay network, which requires feedback from the receiver, was shown to achieve a diversity order of two in Rayleigh fading if the direct source-destination link is available for transmission. These early contributions have sparked a significant interest in cooperative communication techniques which resulted in many new discoveries, e.g. [7]-[15].

### A. Background and Related Work

In practice, half-duplex relays may be preferred as they are easier to implement than full-duplex relays. However, half-duplexing suffers from a multiplexing gain loss compared to full duplexing. To compensate for this loss, existing protocols for the wireless three-node network with a half-duplex relay exploit the direct source-destination link to achieve a throughput gain or a diversity gain over non-relay aided transmission, e.g. [3]-[11]. In practice, because of the typically large distance between source and destination, the direct source-destination link may be very weak and the gains may manifest themselves only at very high signal-to-noise ratios (SNRs). However, if a source-destination link is not available, much of the gains obtained by half-duplex relaying disappear. There are two reasons for this. First, in most of the existing literature, e.g. [3]-[11], the schedule of when the source transmits and when the relay transmits is a priori fixed. Typically, the relay receives a codeword from the source in one time slot and forwards some information about the received codeword to the destination in the next time slot. We refer to this approach in the following as “conventional relaying”. Second, even if the relay has channel state information at the transmitter (CSIT), it does not exploit this information for rate adaptation, see e.g. [12]. In this paper, we propose relaying protocols that select the transmitting node based on the quality of the source-relay and the relay-destination links, i.e., the schedule of transmission is not a priori fixed. For this to be possible, the relays have to be equipped with buffers for data storage, the

node performing the selection of the transmitting node requires the channel state information (CSI) of both involved links, and feedback of a few bits of information from the node performing the selection to the transmitting node is necessary. Furthermore, we assume that if the relay has CSIT, it exploits this knowledge to adapt the transmission rate over the relay-destination channel.

Relays with buffers have been considered in the literature before [16]-[20]. In [16], the buffer at the relay is used to enable the relay to receive for a fixed number of time slots before retransmitting the received information in a fixed number of time slots. In [17], relay selection is considered and buffers enable the selection of the relay with the best source-relay channel for reception and the best relay-destination channel for transmission. However, in both [16] and [17], the schedule of when the source transmits and when the relays transmit is a priori fixed. Thus, these schemes do not achieve a diversity gain compared to conventional relaying. Buffer-aided relaying schemes, where the schedule of when the source transmits and when a relay transmits is not a priori fixed, are considered in [18]-[20]. In [18], the authors propose a protocol for relay selection in a network employing multiple mobile relays with buffers. The protocol operates in one of the following three modes: 1) If there are relay-destination links whose SNR is sufficiently high for successful transmission and the corresponding relays have packets in their buffers, a single relay is chosen to transmit to the destination; else 2) if there are source-relay links with sufficiently high SNR, the source is selected for transmission; else 3) none of the nodes transmits. Furthermore, [19] considers a diamond cooperative network with two relays and buffering at the relays is used only when: 1) The instantaneous SNRs of both source-relay links are smaller than some predefined threshold while the instantaneous SNR of at least one of the relay-destination links is larger than the threshold, or 2) the instantaneous SNRs of both relay-destination links are smaller than the threshold while the instantaneous SNR of at least one of the source-relay links is larger than the threshold. Moreover, the authors in [20] introduce a relay selection scheme for a network employing multiple relays with buffers. In this scheme, the schedule of when a relay receives and transmits depends on the number of packets in the relay's buffer and the instantaneous SNRs of the source-relay and relay-destination links. Although the protocols proposed in [18]-[20] yield a throughput gain over conventional relaying, they were derived based on heuristics, and are thus generally not optimal as far as throughput maximization and/or outage probability minimization are concerned. Consequently, these protocols do not fully exploit the degrees of freedom offered by relays with buffers.

For the case of adaptive rate transmission, the maximum achievable throughput of the simple three-node relay network employing a half-duplex decode-and-forward relay with a buffer was recently derived

in [21]. Thereby, both the source and the relay were assumed to adjust their transmission rate such that outages are avoided. However, adjusting the rate of transmission is not possible if CSIT is not available and/or only one modulation/coding scheme is implemented. In these cases, the protocol proposed in [21] is not applicable. Some preliminary results on buffer-aided relaying for fixed rate transmission have been presented in [22] and independently in [23]. However, although [22], [23] demonstrate that the simple three-node network with one buffer-aided relay and without direct source-destination link can achieve a diversity order of two in Rayleigh fading, the protocols adopted in [22], [23] are suboptimal. Specifically, the protocol in [22] employs a suboptimal decision function for link selection, and the protocol in [23] does not take into account the average link SNRs for link selection, which may lead to low throughputs for non-identical average link SNRs. On the other hand, to the best of our knowledge, relay protocols for the case when the source transmits with fixed rate but the relay can adjust its rate to the channel conditions have not been proposed before.

### B. Contributions

In this paper, we consider the simple three-node relay network with a half-duplex decode-and-forward relay, which is equipped with a buffer, and assume that the direct source-destination link is not available for transmission. We assume that both the source-relay and the relay-destination links are affected by fading. Depending on the availability of CSIT at the transmitting nodes (or their capability of using more than one modulation/coding scheme), we consider two different modes of transmission for the relay network: *Fixed rate transmission* and *mixed rate transmission*. In both modes of transmission, each codeword spans one time slot. In fixed rate transmission, the destination acquires full CSI of both links and selects the link used for transmission, whereas the node selected for transmission (source or relay) does not have CSIT and transmits with fixed rate. In contrast, in mixed rate transmission, the relay obtains full CSI of both involved links and performs link selection. If the relay is selected for transmission, it exploits its CSIT knowledge and transmits with variable rate to avoid outages. However, the source still transmits with fixed rate to avoid the need for CSIT acquisition.

To explore the performance limits of the proposed fixed rate and mixed rate transmission schemes, we consider first transmission without delay constraints and derive the corresponding optimal buffer-aided relaying protocols. Since in practice it is desirable to limit the transmission delay, we also introduce modified buffer-aided relaying protocols for delay constrained transmission. In particular, we make the following main contributions:

- For fixed rate and mixed rate transmission without delay constraints, we derive the optimal buffer-aided relaying protocols which maximize the achievable throughput of the considered three-node relay network employing a half-duplex relay with a buffer of infinite size.
- For fixed rate transmission, we show that in Rayleigh fading the optimal buffer-aided relaying protocol with adaptive link selection achieves a diversity gain of two and a diversity-multiplexing tradeoff of  $DM(r) = 2(1 - 2r)$ , where  $r$  denotes the multiplexing gain.
- For mixed rate transmission, we show that a multiplexing gain of one can be achieved with buffer-aided relaying with and without adaptive link selection implying that there is no multiplexing gain loss compared to ideal full-duplex relaying.
- For fixed rate and mixed rate transmission with delay constraints, in order to control the average delay, we introduce appropriate modifications to the buffer-aided relaying protocols for the delay unconstrained case. Surprisingly, for fixed rate transmission, the full diversity gain is preserved as long as the tolerable average delay exceeds three time slots. For mixed rate transmission with an average delay of  $E\{T\}$  time slots, a multiplexing gain of  $r = 1 - 1/(2E\{T\})$  is achieved.

### C. Organization

The remainder of this paper is organized as follows. In Section II, the system model of the considered three-node relay network is presented. In Sections III and IV, we introduce the proposed buffer-aided relaying protocols for delay unconstrained and delay constrained fixed rate transmission, respectively. Protocols for delay unconstrained and delay constrained mixed rate transmission are proposed and analyzed in Section V. The derived analytical results and relay protocols are verified and illustrated with numerical examples in Section VI, and some conclusions are drawn in Section VII.

## II. SYSTEM MODEL AND BENCHMARK SCHEMES

We consider a three-node wireless network comprising a source  $\mathcal{S}$ , a half-duplex decode-and-forward relay  $\mathcal{R}$ , and a destination  $\mathcal{D}$ , cf. Fig. 1. The source can communicate with the destination only through the relay, i.e., there is no direct  $\mathcal{S}$ - $\mathcal{D}$  link. The source sends codewords to the relay, which decodes these codewords, possibly stores the decoded information in its buffer, and eventually sends it to the destination. We assume that time is divided into slots of equal lengths and every codeword spans one time slot. Throughout this paper, we assume that the source node has always data to transmit. Hence, the total number of time slots, denoted by  $N$ , satisfy  $N \rightarrow \infty$ . Furthermore, unless specified otherwise,

we assume that the buffer at the relay is not limited in size. The case of limited buffer size will be investigated in Sections IV and V-D.

### A. Channel Model

In the  $i$ th time slot, the transmit powers of source and relay are denoted by  $\mathcal{P}_S(i)$  and  $\mathcal{P}_R(i)$ , respectively, and the instantaneous squared channel gains of the  $\mathcal{S}$ - $\mathcal{R}$  and  $\mathcal{R}$ - $\mathcal{D}$  links are denoted by  $h_S(i)$  and  $h_R(i)$ , respectively.  $h_S(i)$  and  $h_R(i)$  are modeled as mutually independent, non-negative, stationary, and ergodic random processes with expected values  $\bar{\Omega}_S \triangleq E\{h_S(i)\}$  and  $\bar{\Omega}_R \triangleq E\{h_R(i)\}$ , where  $E\{\cdot\}$  denotes expectation. We assume that the channel gains are constant during one time slot but change from one time slot to the next due to e.g. the mobility of the involved nodes and/or frequency hopping. We note that for most results derived in this paper, we only require  $h_S(i)$  and  $h_R(i)$  to be not fully temporally correlated, respectively. However, in some cases, we will assume that  $h_S(i)$  and  $h_R(i)$  are temporally uncorrelated, respectively, to facilitate the analysis.

The instantaneous SNRs of the  $\mathcal{S}$ - $\mathcal{R}$  and  $\mathcal{R}$ - $\mathcal{D}$  channels in the  $i$ th time slot are given by  $s(i) \triangleq \gamma_S(i)h_S(i)$  and  $r(i) \triangleq \gamma_R(i)h_R(i)$ , respectively. Here,  $\gamma_S(i) \triangleq \mathcal{P}_S(i)/\sigma_{n_R}^2$  and  $\gamma_R(i) \triangleq \mathcal{P}_R(i)/\sigma_{n_D}^2$  denote the average transmit SNRs of the source and the relay, respectively, and  $\sigma_{n_R}^2$  and  $\sigma_{n_D}^2$  are the variances of the additive white Gaussian noise (AWGN) at the relay and the destination, respectively. The average link SNRs are denoted by  $\Omega_S \triangleq E\{s(i)\}$  and  $\Omega_R \triangleq E\{r(i)\}$ .

Furthermore, for concreteness, we specialize some of the derived results to Rayleigh fading. In this case, the probability density functions (pdfs) of  $s(i)$  and  $r(i)$  are given by  $f_s(s) = e^{-s/\Omega_S}/\Omega_S$  and  $f_r(r) = e^{-r/\Omega_R}/\Omega_R$ , respectively. Similarly, the pdfs of  $h_S(i)$  and  $h_R(i)$  are given by  $f_{h_S}(h_S) = e^{-h_S/\bar{\Omega}_S}/\bar{\Omega}_S$  and  $f_{h_R}(h_R) = e^{-h_R/\bar{\Omega}_R}/\bar{\Omega}_R$ , respectively.

### B. Link Adaptive Transmission Protocol

In the proposed link adaptive transmission protocol, in a given time slot  $i$ , the relay or the destination node acquire knowledge of the instantaneous SNRs  $s(i)$  and  $r(i)$ , and based on this knowledge make a decision whether the source or the relay should transmit. Since the noise variances  $\sigma_{n_R}^2$  and  $\sigma_{n_D}^2$  are constant, we assume that they are known at both the relay and the destination. Thus, the main challenge for buffer-aided adaptive link selection lies in the acquisition of  $h_S(i)$  and  $h_R(i)$ . Depending on whether the relay or the destination acquire  $h_S(i)$  and  $h_R(i)$ , we consider two different transmission modes.

1) *Fixed Rate Transmission:* In the first case, the destination acquires the CSI of both links. Thereby, the destination estimates first the product  $h_S(i)h_R(i)$  based on pilots emitted by the source and forwarded by the relay, and subsequently  $h_R(i)$  based on pilots emitted by the relay. An estimate for  $h_S(i)$  can then be obtained at the destination based on the estimates for  $h_S(i)h_R(i)$  and  $h_R(i)$ . The relay also uses the pilots emitted by the source to estimate  $h_S(i)$ , which is required for coherent detection at the relay. However, source and relay do not have CSIT, i.e., source and relay do not know  $h_S(i)$  and  $h_R(i)$ , respectively. Therefore, both nodes can transmit only with predetermined fixed rates  $S_0$  and  $R_0$ , respectively, and cannot perform power allocation, i.e., the transmit powers are a priori fixed as  $\mathcal{P}_S(i) = \mathcal{P}_S$  and  $\mathcal{P}_R(i) = \mathcal{P}_R$ ,  $\forall i$ . In any given time slot, the destination decides which node is selected for transmission based on the buffer-aided protocol proposed in Section III.

2) *Mixed Rate Transmission:* In the second case, the relay decides which node is selected for transmission. To this end, the relay acquires the CSI of both links based on pilot symbols emitted by source and relay and the assumption of channel reciprocity. To enable coherent detection, the destination also estimates  $h_R(i)$  based on pilots emitted by the relay. Since the relay knows  $h_R(i)$ , it can adjust its transmission rate and transmit power  $\mathcal{P}_R(i)$  to avoid outages on the  $\mathcal{R}-\mathcal{D}$  link. However, the source still has to transmit with fixed rate  $S_0$  and fixed power  $\mathcal{P}_S$  as it does not know the CSI of the  $\mathcal{S}-\mathcal{R}$  channel.

For both modes of transmission, since the node selecting the transmitting node has the CSI of both links and also knows  $S_0$  and  $R_0$ , it can determine a priori whether  $S_0$  and  $R_0$  will exceed the channel capacity of the  $\mathcal{S}-\mathcal{R}$  and the  $\mathcal{R}-\mathcal{D}$  links, respectively. Therefore, if this happens, i.e., if the source (relay) is selected to transmit and the  $\mathcal{S}-\mathcal{R}$  ( $\mathcal{R}-\mathcal{D}$ ) channel is in outage, the node making the selection informs the source (relay) to remain silent and an outage event occurs. Once the decision regarding the transmitting node has been made, the node making this decision broadcasts it to the remaining nodes and transmission in time slot  $i$  begins. The optimization of the buffer-aided link selection protocol for fixed and mixed rate transmission will be discussed in Sections III and V, respectively.

*Remark 1:* We note that in terms of CSI requirements and feedback overhead, the buffer-aided link selection protocols proposed in this paper are very similar to existing relaying protocols, such as the opportunistic protocol proposed in [12]. Namely, the protocol proposed in [12] requires the relays to acquire the instantaneous CSI of the  $\mathcal{S}-\mathcal{R}$  and  $\mathcal{R}-\mathcal{D}$  links. Furthermore, a few bit of information are fed back from the relays to both the source and the destination.



### C. Queue at the Relay

Crucial for derivation of the proposed link selection protocols is a clear understanding of the dynamics of the queue in the buffer of the relay. In the following, for convenience, we normalize the number of bits transmitted in one time slot to the number of symbols per time slot. Thus, throughout the remainder of this paper, when we refer to the number of bits, we mean the number of bits normalized by the number of symbols in a codeword.

If the source is selected for transmission in time slot  $i$  and an outage does not occur, i.e.,  $\log_2(1 + s(i)) \geq S_0$ , it transmits with rate  $S_{\mathcal{SR}}(i) = S_0$ . Hence, the relay receives  $S_0$  data bits from the source and appends them to the queue in its buffer. The number of bits in the buffer of the relay at the end of the  $i$ -th time slot is denoted by  $Q(i)$  and given by

$$Q(i) = Q(i-1) + S_0. \quad (1)$$

If the source is selected for transmission but the  $\mathcal{S}$ - $\mathcal{R}$  link is in outage, i.e.,  $\log_2(1 + s(i)) < S_0$ , the source remains silent, i.e.,  $S_{\mathcal{SR}}(i) = 0$ , and the queue in the buffer remains unchanged, i.e.,  $Q(i) = Q(i-1)$ .

For fixed rate transmission, if the relay is selected for transmission in time slot  $i$  and transmits with rate  $R_0$ , an outage does not occur if  $\log_2(1 + r(i)) \geq R_0$ . In this case, the number of bits transmitted by the relay is given by

$$R_{\mathcal{RD}}(i) = \min\{R_0, Q(i-1)\}, \quad (2)$$

where we take into account that the maximum number of bits that can be send by the relay is also limited by the number of bits in its buffer. The number of data bits remaining in the buffer at the end of time slot  $i$  is given by

$$Q(i) = Q(i-1) - R_{\mathcal{RD}}(i), \quad (3)$$

which is always non-negative because of (2). If the relay is selected for transmission in time slot  $i$  but an outage occurs, i.e.,  $\log_2(1 + r(i)) < R_0$ , the relay remains silent, i.e.,  $R_{\mathcal{RD}}(i) = 0$ , while the queue in the buffer remains unchanged, i.e.,  $Q(i) = Q(i-1)$ .

For mixed rate transmission, the relay is able to adapt its rate to the capacity of the  $\mathcal{R}$ - $\mathcal{D}$  channel,  $\log_2(1 + r(i))$ , and outages are avoided. If the relay is selected for transmission in time slot  $i$ , the number of bits transmitted by the relay is given by

$$R_{\mathcal{RD}}(i) = \min\{\log_2(1 + r(i)), Q(i-1)\}. \quad (4)$$



The number of data bits remaining in the buffer at the end of time slot  $i$  is still given by (3) where  $R_{\mathcal{RD}}(i)$  is now given by (4).

Furthermore, because of the half-duplex constraint, for both fixed and mixed rate transmission, we have  $R_{\mathcal{RD}}(i) = 0$  and  $S_{\mathcal{SR}}(i) = 0$  if source and relay are selected for transmission in time slot  $i$ , respectively.

#### D. Link Outages and Indicator Variables

For future reference, we introduce the binary link outage indicator variables  $O_S(i) \in \{0, 1\}$  and  $O_R(i) \in \{0, 1\}$  defined as

$$O_S(i) \triangleq \begin{cases} 0 & \text{if } s(i) < 2^{S_0} - 1 \\ 1 & \text{if } s(i) \geq 2^{S_0} - 1 \end{cases} \quad (5)$$

and

$$O_R(i) \triangleq \begin{cases} 0 & \text{if } r(i) < 2^{R_0} - 1 \\ 1 & \text{if } r(i) \geq 2^{R_0} - 1 \end{cases} \quad (6)$$

respectively. In other words,  $O_S(i) = 0$  indicates that for transmission with rate  $S_0$ , the  $\mathcal{S}$ - $\mathcal{R}$  link is in outage, i.e.,  $\log_2(1 + s(i)) < S_0$ , and  $O_S(i) = 1$  indicates that the transmission over the  $\mathcal{S}$ - $\mathcal{R}$  channel will be successful. Similarly,  $O_R(i) = 0$  indicates that for transmission with rate  $R_0$ , the  $\mathcal{R}$ - $\mathcal{D}$  link is in outage, i.e.,  $\log_2(1 + r(i)) < R_0$ , and  $O_R(i) = 1$  means that an outage will not occur. Furthermore, we denote the outage probabilities of the  $\mathcal{S}$ - $\mathcal{R}$  and  $\mathcal{R}$ - $\mathcal{D}$  channels as  $P_S$  and  $P_R$ , respectively. These probabilities are defined as

$$P_S \triangleq \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N (1 - O_S(i)) = \Pr\{s(i) < 2^{S_0} - 1\} \quad (7)$$

and

$$P_R \triangleq \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N (1 - O_R(i)) = \Pr\{r(i) < 2^{R_0} - 1\}, \quad (8)$$

respectively.

#### E. Performance Metrics

In this paper, we adopt the throughput and the outage probability as performance metrics.

Assuming the source has always data to transmit, for both fixed and mixed rate transmission, the average number of bits that arrive at the destination per time slot is given by

$$\tau = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N R_{\mathcal{RD}}(i), \quad (9)$$

i.e.,  $\tau$  is the throughput of the considered communication system.

The outage probability is defined as the probability that the instantaneous channel capacity is unable to support some predetermined fixed transmission rate. In the considered system, an outage does not cause information loss since the node performing link selection knows in advance whether or not the selected link can support the chosen transmission rate and inform the transmitting node accordingly. Nevertheless, outages still affect the achievable throughput negatively. In fact, the outage probability can be interpreted as the *fraction of the throughput lost due to outages*. Thus, denoting the maximum throughput of a system in the absence of outages by  $\tau_0$  and the throughput in the presence of outages by  $\tau$ , the outage probability,  $F_{\text{out}}$ , can be expressed as

$$F_{\text{out}} = 1 - \frac{\tau}{\tau_0}. \quad (10)$$

Note that maximizing the throughput is equivalent to minimizing the outage probability.

#### F. Performance Benchmarks for Fixed Rate Transmission

For fixed rate transmission, two conventional relaying schemes serve as performance benchmarks for the proposed buffer-aided relaying scheme with adaptive link selection. In contrast to the proposed scheme, the benchmark schemes employ a predetermined schedule for when source and relay transmit which is independent of the instantaneous link SNRs.

In the first scheme, referred to as Conventional Relaying 1 (see also [16]), the source transmits in the first  $\xi N$  time slots, where  $0 < \xi < 1$  and each codeword spans one time slot. The relay tries to decode these codewords and, if the decoding is successful, it stores the corresponding information bits in its buffer. In the following  $(1 - \xi)N$  time slots, the relay transmits the stored information bits to the destination, transmitting one codeword per time slot. Assuming that for the benchmark schemes source and relay transmit codewords having the same rate, i.e.,  $S_0 = R_0$ , the throughput of Conventional Relaying 1 is obtained as

$$\begin{aligned} \tau_{\text{conv},1}^{\text{fixed}} &= \lim_{N \rightarrow \infty} \frac{1}{N} \min \left\{ \sum_{i=1}^{\xi N} R_0 O_S(i), \sum_{i=\xi N+1}^N R_0 O_R(i) \right\} \\ &= R_0 \min \{ \xi(1 - P_S), (1 - \xi)(1 - P_R) \}. \end{aligned} \quad (11)$$

The throughput is maximized if  $\xi(1 - P_S) = (1 - \xi)(1 - P_R)$  holds or equivalently if  $\xi = (1 - P_R)/(2 - P_S - P_R)$ . Inserting  $\xi$  into (11) we obtain the maximized throughput as

$$\tau_{\text{conv},1}^{\text{fixed}} = R_0 \frac{(1 - P_S)(1 - P_R)}{2 - P_S - P_R}. \quad (12)$$

The maximum throughput in the absence of outages is  $\tau_0 = R_0/2$ , hence using (10), the corresponding outage probability is obtained as

$$F_{\text{out,conv},1}^{\text{fixed}} = 1 - 2 \frac{(1 - P_S)(1 - P_R)}{2 - P_S - P_R}. \quad (13)$$

In the second scheme, referred to as Conventional Relaying 2, in the first time slot, the source transmits one codeword and the relay receives and tries to decode the codeword. If the decoding is successful, in the second time slot, the relay retransmits the information to the destination, otherwise it remains silent. The throughput of Conventional Relaying 2 is obtained as

$$\tau_{\text{conv},2}^{\text{fixed}} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^{N/2} R_0 O_S(2i - 1) O_R(2i) = \frac{R_0}{2} (1 - P_S)(1 - P_R). \quad (14)$$

Based on (10) the corresponding outage probability is given by

$$F_{\text{out,conv},2}^{\text{fixed}} = 1 - (1 - P_S)(1 - P_R). \quad (15)$$

We note that  $\tau_{\text{conv},1}^{\text{fixed}} \geq \tau_{\text{conv},2}^{\text{fixed}}$  ( $F_{\text{out,conv},1}^{\text{fixed}} \leq F_{\text{out,conv},2}^{\text{fixed}}$ ) always holds. However, in order for Conventional Relaying 1 to realize this gain, an infinite delay is required, whereas Conventional Relaying 2 requires a delay of only one time slot.

For the special case of Rayleigh fading, we obtain from (7) and (8)  $P_S = 1 - e^{-\frac{2R_0-1}{\Omega_S}}$  and  $P_R = 1 - e^{-\frac{2R_0-1}{\Omega_R}}$ , respectively. The corresponding throughputs and outage probabilities for Conventional Relaying 1 and 2 can be obtained by applying these results in (12)-(15). In particular, in the high SNR regime, when  $\gamma_S = \gamma_R = \gamma \rightarrow \infty$ , we obtain  $\tau_{\text{conv},1}^{\text{fixed}} \rightarrow R_0/2$ ,  $\tau_{\text{conv},2}^{\text{fixed}} \rightarrow R_0/2$ , and

$$F_{\text{out,conv},1}^{\text{fixed}} \rightarrow \frac{2^{R_0} - 1}{2} \frac{\bar{\Omega}_S + \bar{\Omega}_R}{\bar{\Omega}_S \bar{\Omega}_R} \frac{1}{\gamma}, \quad \text{as } \gamma \rightarrow \infty \quad (16)$$

$$F_{\text{out,conv},2}^{\text{fixed}} \rightarrow (2^{R_0} - 1) \frac{\bar{\Omega}_S + \bar{\Omega}_R}{\bar{\Omega}_S \bar{\Omega}_R} \frac{1}{\gamma}, \quad \text{as } \gamma \rightarrow \infty \quad (17)$$

Hence, for fixed rate transmission, the diversity gain of Conventional Relaying 1 and 2 is one as expected.

### G. Performance Benchmarks for Mixed Rate Transmission

We also provide two performance benchmarks with a priori fixed link selection schedule for mixed rate transmission. The two benchmark protocols are analogous to the corresponding protocols in the fixed rate case. Thus, for Conventional Relaying 1, the source transmits in the first  $\xi N$  time slots with fixed rate  $S_0$  and the relay transmits in the remaining  $(1 - \xi)N$  time slots with rate  $R(i) = \log_2(1 + r(i))$ . Thus, the throughput is given by

$$\begin{aligned}\tau_{\text{conv},1}^{\text{mixed}} &= \lim_{N \rightarrow \infty} \frac{1}{N} \min \left\{ \sum_{i=1}^{\xi N} S_0 O_S(i) , \sum_{i=\xi N+1}^N \log_2(1 + r(i)) \right\} \\ &= \min \{ \xi(1 - P_S)S_0 , (1 - \xi)E\{\log_2(1 + r(i))\} \}.\end{aligned}\quad (18)$$

The throughput is maximized if  $\xi$  satisfies

$$\xi S_0(1 - P_S) = (1 - \xi)E\{\log_2(1 + r(i))\} . \quad (19)$$

From (19), we obtain  $\xi$  as

$$\xi = \frac{E\{\log_2(1 + r(i))\}}{S_0(1 - P_S) + E\{\log_2(1 + r(i))\}}. \quad (20)$$

Inserting  $\xi$  in the left (or right) hand side of (19), leads to the throughput of mixed rate transmission under the Conventional Relaying 1 protocol

$$\tau_{\text{conv},1}^{\text{mixed}} = \frac{S_0(1 - P_S)E\{\log_2(1 + r(i))\}}{S_0(1 - P_S) + E\{\log_2(1 + r(i))\}}. \quad (21)$$

Assuming Rayleigh fading links  $E\{\log_2(1 + r(i))\}$  is obtained as

$$E\{\log_2(1 + r(i))\} = \frac{e^{1/\Omega_R}}{\ln(2)} E_1\left(\frac{1}{\Omega_R}\right) \quad (22)$$

for fixed transmit powers, where  $E_1(x) = \int_x^\infty e^{-t}/t dt$ ,  $x > 0$ , denotes the exponential integral function.

If adaptive power allocation is employed,  $E\{\log_2(1 + r(i))\}$  becomes

$$E\{\log_2(1 + r(i))\} = \frac{1}{\ln(2)} E_1\left(\frac{\lambda_c}{\Omega_R}\right), \quad (23)$$

where  $\lambda_c$  is found from the power constraint

$$(1 - P_S)\gamma_S + \int_{\lambda_c}^\infty \left(\frac{1}{\lambda_c} - \frac{1}{h_R}\right) f_{h_R}(h_R) dh_R = 2\Gamma. \quad (24)$$

Here,  $\Gamma$  denotes the average transmit power in one time slot. In the high SNR regime, where  $\gamma_S = \gamma_R = \gamma \rightarrow \infty$ ,  $E\{\log_2(1 + r(i))\} \gg S_0(1 - P_S)$  holds. Thus, the throughput in (21) converges to

$$\tau_{\text{conv},1}^{\text{mixed}} \rightarrow S_0, \quad \text{as } \gamma \rightarrow \infty, \quad (25)$$

which leads to the interesting conclusion that mixed rate transmission achieves a multiplexing rate of one even if suboptimal conventional relaying is used.

For Conventional Relaying 2, the performance of mixed rate transmission is identical to that of fixed rate transmission. Since the relay does not employ a buffer for Conventional Relaying 2, even with mixed rate transmission, the relay can only transmit successfully all of the received information if  $S_0 \leq \log_2(1 + r(i))$  and has to remain silent otherwise.

### III. FIXED RATE TRANSMISSION WITHOUT DELAY CONSTRAINTS

In this section, we investigate buffer-aided relaying with adaptive link selection for fixed rate transmission without delay constraints, i.e., the transmission rates of the source and the relay are fixed. We derive the optimal link selection protocol and analyze the corresponding throughput and outage probability. The obtained results constitute performance upper bounds for fixed rate transmission with delay constraints, which will be considered in Section IV.

#### A. Problem Formulation

First, we introduce the binary link selection variable  $d_i \in \{0, 1\}$ . Here,  $d_i = 1$  indicates that the  $\mathcal{R}$ - $\mathcal{D}$  link is selected for transmission in time slot  $i$ , i.e., the relay transmits and the destination receives. Similarly, if  $d_i = 0$ , the  $\mathcal{S}$ - $\mathcal{R}$  link is selected for transmission in time slot  $i$ , i.e., the source transmits and the relay receives.

Based on the definitions of  $O_S(i)$ ,  $O_R(i)$ , and  $d_i$ , the number of bits sent from the source to the relay and from the relay to the destination in time slot  $i$  can be written in compact form as

$$S_{S\mathcal{R}}(i) = (1 - d_i)O_S(i)S_0 \quad (26)$$

and

$$R_{\mathcal{R}\mathcal{D}}(i) = d_i O_R(i) \min\{R_0, Q(i - 1)\}, \quad (27)$$

respectively. Consequently, the throughput in (9) can be rewritten as

$$\tau = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N d_i O_R(i) \min\{R_0, Q(i - 1)\}. \quad (28)$$

In the following, we maximize the throughput by optimizing the link selection variable  $d_i$ , which represents the only degree of freedom in the considered problem. In particular, as already mentioned in Section II-B, since both transmitting nodes do not have the CSI of their respective transmit channels, power allocation is not possible and we assume fixed transmit powers  $\mathcal{P}_S(i) = \mathcal{P}_S$  and  $\mathcal{P}_R(i) = \mathcal{P}_R, \forall i$ .

### B. Throughput Maximization

Let us first define the average arrival rate of bits per slot into the queue of the buffer,  $A$ , and the average departure rate of bits per slot out of the queue of the buffer,  $D$ , as [24]

$$A \triangleq \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N (1 - d_i) O_S(i) S_0 \quad (29)$$

and

$$D \triangleq \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N d_i O_R(i) \min\{R_0, Q(i-1)\}, \quad (30)$$

respectively. We note that the departure rate of the queue is equal to the throughput. The queue is said to be an absorbing queue if  $A > D = \tau$ , in which case a fraction of the information sent by the source is trapped in the buffer and can never be extracted from it. The following theorem provides a useful condition for the optimal policy which maximizes the throughput.

*Theorem 1:* The link selection policy that maximizes the throughput of the considered buffer-aided relaying system can be found in the set of link selection policies that satisfy

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N (1 - d_i) O_S(i) S_0 = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N d_i O_R(i) R_0, \quad (31)$$

and the throughput is given by the right (or left) hand side of (31). If (31) holds, the queue is non-absorbing but is at the edge of absorption, i.e., a small increase of the arrival rate will lead to an absorbing queue.

*Proof:* Please refer to Appendix A. ■

Note that the  $\min(\cdot)$  function in (28) is absent in the throughput in (31), which is crucial for finding a tractable analytical expression for the optimal link selection policy. In particular, when (31) holds, the event that the relay is selected for transmission and its buffer is empty, i.e.,  $Q(i-1) < R_0$ , has a negligible effect on the throughput as  $N \rightarrow \infty$ , since the buffer is practically always fully backlogged.

According to Theorem 1, in order to maximize the throughput, we have to search for the optimal policy only in the set of policies that satisfy (31). Therefore, the search for the optimal policy can be formulated as an optimization problem, which for  $N \rightarrow \infty$  has the following form

$$\begin{aligned} \text{Maximize : } & \frac{1}{N} \sum_{i=1}^N d_i O_R(i) R_0 \\ \text{Subject to : } & \text{C1 : } \frac{1}{N} \sum_{i=1}^N (1 - d_i) O_S(i) S_0 = \frac{1}{N} \sum_{i=1}^N d_i O_R(i) R_0 \\ & \text{C2 : } \frac{1}{N} d_i (1 - d_i) = 0, \quad \forall i \end{aligned} \quad (32)$$

where constraint C1 ensures that the search for the optimal policy is conducted only among those policies that satisfy (31) and C2 ensures that  $d_i \in \{0, 1\}$ .

Before we solve problem (32), we note that, as will be shown in the following, the optimal link selection policy may require a coin flip. For this purpose, we introduce the set of possible outcomes of the coin flip,  $\mathcal{C} \in \{0, 1\}$ , and denote the probabilities of the outcomes by  $P_C = \Pr\{\mathcal{C} = 1\}$  and  $\Pr\{\mathcal{C} = 0\} = 1 - P_C$ , respectively. Now, we are ready to provide the solution of (32), which constitutes the optimal link selection policy maximizing the throughput. This is conveyed in the following theorem.

*Theorem 2:* For the optimal link selection policy maximizing the throughput of the considered buffer-aided relaying system for fixed rate transmission, three mutually exclusive cases can be distinguished depending on the values of  $P_S$  and  $P_R$ :

**Case 1:**

$$P_S \leq \frac{S_0}{S_0 + R_0(1 - P_R)} \quad \text{AND} \quad P_R \leq \frac{R_0}{R_0 + S_0(1 - P_S)}. \quad (33)$$

In this case, the optimal link selection policy is given by

$$d_i = \begin{cases} 0 & \text{if } O_S(i) = 1 \text{ AND } O_R(i) = 0 \\ 1 & \text{if } O_S(i) = 0 \text{ AND } O_R(i) = 1 \\ 0 & \text{if } O_S(i) = 1 \text{ AND } O_R(i) = 1 \text{ AND } \mathcal{C} = 0 \\ 1 & \text{if } O_S(i) = 1 \text{ AND } O_R(i) = 1 \text{ AND } \mathcal{C} = 1 \\ \varepsilon & \text{if } O_S(i) = 0 \text{ AND } O_R(i) = 0 \end{cases} \quad (34)$$

where  $\varepsilon$  can be set to 0 or 1 as neither the source nor the relay will transmit because both links are in outage. On the other hand, if both links are not in outage, i.e.,  $O_S(i) = 1$  and  $O_R(i) = 1$ , the coin flip decides which node transmits and the probability of  $\mathcal{C} = 1$  is given by

$$P_C = \frac{S_0(1 - P_S) - (1 - P_R)P_S R_0}{(1 - P_S)(1 - P_R)(S_0 + R_0)}. \quad (35)$$

Based on (34), the maximum throughput is obtained as

$$\tau = \frac{S_0 R_0}{S_0 + R_0} (1 - P_S P_R). \quad (36)$$

**Case 2:**

$$P_R > \frac{R_0}{R_0 + S_0(1 - P_S)} \quad (37)$$



In this case, the optimal link selection policy is characterized by

$$d_i = \begin{cases} 0 & \text{if } O_S(i) = 1 \text{ AND } O_R(i) = 0 \text{ AND } \mathcal{C} = 0 \\ 1 & \text{if } O_S(i) = 1 \text{ AND } O_R(i) = 0 \text{ AND } \mathcal{C} = 1 \\ 1 & \text{if } O_S(i) = 0 \text{ AND } O_R(i) = 1 \\ 1 & \text{if } O_S(i) = 1 \text{ AND } O_R(i) = 1 \\ \varepsilon & \text{if } O_S(i) = 0 \text{ AND } O_R(i) = 0 \end{cases} \quad (38)$$

The probability of outcome  $\mathcal{C} = 1$  of the coin flip is given by

$$P_C = \frac{S_0(1 - P_S)P_R - (1 - P_R)R_0}{(1 - P_S)P_RS_0}, \quad (39)$$

and the maximum throughput can be obtained as

$$\tau = R_0(1 - P_R). \quad (40)$$

**Case 3:**

$$P_S > \frac{S_0}{S_0 + R_0(1 - P_R)}. \quad (41)$$

In this case, the link selection policy that maximizes the throughput is given by

$$d_i = \begin{cases} 0 & \text{if } O_S(i) = 1 \text{ AND } O_R(i) = 0 \\ 0 & \text{if } O_S(i) = 0 \text{ AND } O_R(i) = 1 \text{ AND } \mathcal{C} = 0 \\ 1 & \text{if } O_S(i) = 0 \text{ AND } O_R(i) = 1 \text{ AND } \mathcal{C} = 1 \\ 0 & \text{if } O_S(i) = 1 \text{ AND } O_R(i) = 1 \\ \varepsilon & \text{if } O_S(i) = 0 \text{ AND } O_R(i) = 0 \end{cases} \quad (42)$$

The probability of  $\mathcal{C} = 1$  is given by

$$P_C = \frac{S_0(1 - P_S)}{R_0(1 - P_R)P_S}, \quad (43)$$

and the maximum throughput is

$$\tau = S_0(1 - P_S). \quad (44)$$

*Proof:* Please refer to Appendix B. ■

*Remark 2:* We note that in the second line of (38), we set  $d_i = 1$  although the  $\mathcal{R}$ - $\mathcal{D}$  link is in outage ( $O_R(i) = 0$ ) while the  $\mathcal{S}$ - $\mathcal{R}$  link is not in outage ( $O_S(i) = 1$ ). In other words, in this case, neither node transmits although the source node could successfully transmit. However, if the source node transmitted in this situation, the queue at the relay would become an absorbing queue. Similarly, in the second line

of (42), we set  $d_i = 0$  although the  $\mathcal{S}$ - $\mathcal{R}$  link is in outage. Again, neither node transmits in order to ensure that condition (31) is met. However, in this case, the exact same throughput as in (44) can be achieved with a simpler and more practical link selection policy than that in (42). This is addressed in the following lemma.

*Lemma 1:* The throughput achieved by the link selection policy in (42) can also be achieved with the following simpler link selection policy.

If

$$P_S > \frac{S_0}{S_0 + R_0(1 - P_R)}, \quad (45)$$

a link selection policy maximizing the throughput is given by

$$d_i = \begin{cases} 0 & \text{if } O_S(i) = 1 \\ 1 & \text{if } O_S(i) = 0 \end{cases} \quad (46)$$

and the maximum throughput is

$$\tau = S_0(1 - P_S). \quad (47)$$

*Proof:* The policy given by (42) has the same average arrival rate as policy (46) since for both policies the source always transmits when  $O_S(i) = 1$ . Therefore, since for both policies the queue is non-absorbing, by the law of conservation of flow, their throughputs are identical to their arrival rates. Thus, both policies achieve identical throughputs. ■

*Remark 3:* Note that when  $P_R > R_0/(R_0 + S_0(1 - P_S))$  ( $P_S > S_0/(S_0 + R_0(1 - P_R))$ ) holds, the throughput is given by (40) ((44)), which is identical to the maximal throughput that can be obtained in a point-to-point communication between relay and destination (source and relay). Therefore, when  $P_R > R_0/(R_0 + S_0(1 - P_S))$  ( $P_S > S_0/(S_0 + R_0(1 - P_R))$ ) holds, as far as the achievable throughput is concerned, the three-node half-duplex relay channel is equivalent to the two-node  $\mathcal{R}$ - $\mathcal{D}$  ( $\mathcal{S}$ - $\mathcal{R}$ ) channel.

For comparison, we also provide the maximum throughput in the absence of outages  $\tau_0$ . The throughput in the absence of outages,  $\tau_0$ , can be obtained by setting  $O_S(i) = O_R(i) = 1, \forall i$ , which is equivalent to setting  $P_S = P_R = 0$  in Theorem 2. Then, Case 1 in Theorem 2 always holds and the optimal link selection policy is

$$d_i = \begin{cases} 0 & \text{if } \mathcal{C} = 0 \\ 1 & \text{if } \mathcal{C} = 1 \end{cases} \quad (48)$$

where the probability of  $\mathcal{C} = 1$  is given by

$$P_C = \frac{S_0}{S_0 + R_0}. \quad (49)$$

Based on (48), the maximum throughput in the absence of outages is

$$\tau_0 = \frac{S_0 R_0}{S_0 + R_0}. \quad (50)$$

The throughput loss caused by outages can be observed by comparing (36), (40), and (44) with (50).

We now provide the outage probability of the proposed buffer-aided relaying scheme with adaptive link selection.

*Lemma 2:* The outage probability of the system considered in Theorem 2 is given by

$$F_{\text{out}} = \begin{cases} P_R - (1 - P_R)R_0/S_0, & \text{if } P_R > R_0/(R_0 + S_0(1 - P_S)) \\ P_S - (1 - P_S)S_0/R_0, & \text{if } P_S > S_0/(S_0 + R_0(1 - P_R)) \\ P_S P_R, & \text{otherwise} \end{cases} \quad (51)$$

*Proof:* Please refer to Appendix C. ■

*Remark 4:* In the proof of Lemma 2 given in Appendix C, it is shown that an outage event happens when neither the source nor the relay transmit in a time slot, i.e., the number of silent slots is identical to the number of outage events.

In the high SNR regime, when the outage probabilities of both involved links are small, the expressions for the throughput and the outage probability can be simplified to obtain further insight into the performance of buffer-aided relaying. This is addressed in the following lemma.

*Lemma 3:* In the high SNR regime,  $\gamma_S = \gamma_R = \gamma \rightarrow \infty$ , the throughput and the outage probability of the buffer-aided relaying system considered in Theorem 2 converge to

$$\tau \rightarrow \tau_0 = \frac{S_0 R_0}{S_0 + R_0}, \quad \text{as } \gamma \rightarrow \infty \quad (52)$$

$$F_{\text{out}} = P_S P_R \quad (53)$$

*Proof:* In the high SNR regime, we have  $P_S \rightarrow 0$  and  $P_R \rightarrow 0$ . Thus, condition (33) always holds and therefore  $F_{\text{out}}$  is given by (53). Furthermore, as  $P_S \rightarrow 0$  and  $P_R \rightarrow 0$ , (36) simplifies to (52). ■

### C. Performance in Rayleigh Fading

For concreteness, we assume in this subsection that both links of the considered three-node relay system are Rayleigh fading. We examine the diversity order and the diversity-multiplexing trade-off.

*Lemma 4:* For the special case of Rayleigh fading links, the buffer-aided relaying system considered in Theorem 2 achieves a diversity gain of two, i.e., in the high SNR regime, when  $\gamma_S = \gamma_R = \gamma \rightarrow \infty$ , the outage probability,  $F_{\text{out}}$ , decays on a log-log scale with slope  $-2$  as a function of the transmit SNR  $\gamma$ , and is given by

$$F_{\text{out}} \rightarrow \frac{2^{S_0} - 1}{\bar{\Omega}_S} \frac{2^{R_0} - 1}{\bar{\Omega}_R} \frac{1}{\gamma^2}, \quad \text{as } \gamma \rightarrow \infty. \quad (54)$$

Furthermore, the considered buffer-aided relaying system achieves a diversity-multiplexing trade-off,  $DM(r)$ , of

$$DM(r) = 2(1 - 2r), \quad 0 < r < 1/2. \quad (55)$$

*Proof:* Please refer to Appendix D. ■

*Remark 5:* We recall that, for fixed rate transmission, both considered conventional relay schemes without adaptive link selection achieved only a diversity gain of one, cf. (16), (17), despite the fact that Conventional Relaying 1 also entails an infinite delay. Thus, we expect large gains in terms of outage probability of the proposed buffer-aided relaying protocol with adaptive link selection compared to conventional relaying.

The performance of the considered system can be further improved by optimizing the transmission rates  $R_0$  and  $S_0$  based on the channel statistics. For Rayleigh fading with given  $\bar{\Omega}_S$  and  $\bar{\Omega}_R$ , we can optimize  $R_0$  and  $S_0$  for minimization of the outage probability. This is addressed in the following lemma.

*Lemma 5:* Assuming Rayleigh fading, the optimal transmission rates  $S_0$  and  $R_0$  that minimize the outage probability in the high SNR regime, while maintaining a throughput of  $\tau_0$ , are given by  $R_0 = S_0 = 2\tau_0$ .

*Proof:* The throughput in the high SNR regime is given by (52), which can be rewritten as  $R_0 = S_0\tau_0/(S_0 - \tau_0)$ . Inserting this into the asymptotic expression for  $F_{\text{out}}$  in (54) and minimizing it with respect to  $S_0$  yields  $S_0 = R_0 = 2\tau_0$ . ■

*Remark 6:* For Rayleigh fading, although in the low SNR regime, the optimal  $S_0$  and  $R_0$  can be nonidentical, in the high SNR regime, independent of the values of  $\bar{\Omega}_S$  and  $\bar{\Omega}_R$ , the minimum  $F_{\text{out}}$  is obtained for identical transmission rates for both links. Furthermore, in the high SNR regime, when  $\gamma_S = \gamma_R \rightarrow \infty$ , for  $S_0 = R_0$ , the coin flip probability  $P_C$  converges to  $P_C = \Pr\{\mathcal{C} = 1\} = \Pr\{\mathcal{C} = 0\} \rightarrow 1/2$ .

#### IV. FIXED RATE TRANSMISSION WITH DELAY CONSTRAINTS

The protocol proposed in Section III does not impose any constraint on the delay that a transmitted bit experiences. However, in practice, most communication services require delay constraints. Therefore, in this section, we modify the buffer-aided relaying protocol derived in the previous section to account for constraints on the average delay. Furthermore, we analyze the effect of the applied modification on the throughput and the outage probability. Since we found in the previous section that, for high SNR, the source and relay transmission rates should be chosen identical, for simplicity, throughout this section, we assume  $S_0 = R_0$ . Furthermore, to facilitate our analysis, throughout this section, we assume temporally uncorrelated fading.

##### A. Preliminaries

We define the delay of a bit as the time interval from its transmission by the source to its reception at the destination. Thus, assuming that the propagation delays in the  $\mathcal{S}\text{-}\mathcal{R}$  and  $\mathcal{R}\text{-}\mathcal{D}$  links are negligible, the delay of a bit is identical to the time that the bit is held in the buffer. As a consequence, we can use Little's law [25] and express the average delay as

$$E\{T\} = \frac{E\{Q\}}{A}, \quad (56)$$

where  $E\{Q\} = \lim_{N \rightarrow \infty} \sum_{i=1}^N Q(i)/N$  is the average length of the queue in the buffer of the relay and  $A$  is the arrival rate in bits/slot into the queue as defined in (29). Since  $E\{Q\}$  is given in bits and  $A$  is given in bits/slot, the average delay  $E\{T\}$  is given in time slots. From (56), we observe that the delay can be controlled via the queue size.

##### B. Link Selection Protocol for Delay Limited Transmission

As mentioned before, we modify the optimal link selection protocol derived in Section III in order to limit the average delay. However, depending on the targeted average delay, somewhat different modifications are necessary, since it is not possible to achieve any desired delay with one protocol. Hence, three different link selection protocols are introduced in the following proposition.

*Proposition 1:* For fixed rate transmission with delay constraint, depending on the targeted average delay  $E\{T\}$  and the outage probabilities  $P_S$  and  $P_R$ , we propose the following policies:

**Case 1:** If  $P_R < 1/(2 - P_S)$  and the required delay  $E\{T\}$  satisfies

$$E\{T\} > \frac{1}{1 - P_R(2 - P_S)} + \frac{2(1 - P_S)}{1 - P_S P_R(2 - P_S)}, \quad (57)$$

we propose the following link selection variable  $d_i$  to be used:

$$\text{If } Q(i-1) \leq R_0 \text{ and } O_S(i) = 1, \text{ then } d_i = 0, \text{ otherwise } d_i \text{ is given by (34).} \quad (58)$$

**Case 2:** If  $P_R < 1/(2 - P_S)$  and the required delay  $E\{T\}$  satisfies

$$\frac{1}{1 - P_R(2 - P_S)} < E\{T\} \leq \frac{1}{1 - P_R(2 - P_S)} + \frac{2(1 - P_S)}{1 - P_S P_R(2 - P_S)}, \quad (59)$$

we propose the following link selection variable  $d_i$  to be used:

$$\text{If } Q(i-1) = 0 \text{ and } O_S(i) = 1, \text{ then } d_i = 0, \text{ otherwise } d_i \text{ is given by (34).} \quad (60)$$

**Case 3:** If the required delay  $E\{T\}$  satisfies

$$\frac{1}{1 - P_R} < E\{T\} \leq \frac{1}{1 - P_R(2 - P_S)}, \quad (61)$$

we propose the following link selection variable  $d_i$  to be used:

$$\text{If } Q(i-1) = 0 \text{ and } O_S(i) = 1, \text{ then } d_i = 0, \text{ otherwise } d_i \text{ is given by (38).} \quad (62)$$

For each of the proposed link selection variables  $d_i$ , the required delay can be met by adjusting the value of  $P_C = \Pr\{C = 1\}$ , where the minimum and maximum delays are achieved with  $P_C = 1$  and  $P_C = 0$ , respectively.

*Remark 7:* The delay limits given by (57), (59), and (61) arise from the analysis of the proposed protocols with link selection variables (58), (60), and (62), respectively. We will investigate these delay limits in Lemma 7 in Section IV-C and the corresponding proof is provided in Appendix G.

*Remark 8:* We have not proposed a buffer-aided relaying protocol with adaptive link selection that can satisfy a required delay smaller than  $1/(1 - P_R)$ . For such small delays, Conventional Relaying 2 without adaptive link selection can be used.

### C. Throughput and Delay

In the following, we analyze the throughput, the average delay, and the probability of having  $k$  packets in the queue for the modified link selection protocols proposed in Proposition 1 in the previous subsection. The results are summarized in the following theorem.

*Theorem 3:* Consider a buffer-aided relaying system operating in temporally uncorrelated block fading. Let source and relay transmit with rate  $R_0$ , respectively, and let the buffer size at the relay be limited to  $L$  packets each comprised of  $R_0$  bits. Assume that the relay drops newly received packets if the buffer

is full. Then, depending on the adopted link selection protocol, the following cases can be distinguished:

**Case 1:** If the link selection variable  $d_i$  is given by (58), the probability of the buffer having  $k$  packets in its queue,  $\Pr\{Q = kR_0\}$ , is obtained as

$$\Pr\{Q = kR_0\} = \begin{cases} \frac{p^{L-1}(2p+q-1)(P_S-q)}{p^{L-1}(2p(1-q)+q(2-q)-P_S(2-P_S))-(1-p-q)^{L-1}(1-P_S)^2}, & k = 0 \\ \frac{p^{L-1}(2p+q-1)(1-P_S)}{p^{L-1}(2p(1-q)+q(2-q)-P_S(2-P_S))-(1-p-q)^{L-1}(1-P_S)^2}, & k = 1 \\ \frac{p^{L-k}(2p+q-1)(1-P_S)^2(1-p-q)^{k-2}}{p^{L-1}(2p(1-q)+q(2-q)-P_S(2-P_S))-(1-p-q)^{L-1}(1-P_S)^2}, & k = 2 \dots L \end{cases} \quad (63)$$

where  $p$  and  $q$  are given by

$$p = (1 - P_S)(1 - P_R)P_C + P_S(1 - P_R) \quad \text{and} \quad q = P_S P_R. \quad (64)$$

Furthermore, the average queue length,  $E\{Q\}$ , the average delay,  $E\{T\}$ , and throughput,  $\tau$ , are given by

$$E\{Q\} = R_0 \frac{1 - P_S}{2p + q - 1} \times \frac{p^{L-1}((2p+q)^2 - p - q - P_S(3p+q-1)) - (1-P_S)(1-p-q)^{L-1}(L(2p+q-1) + p)}{p^{L-1}(2p(1-q) + (2-q)q - (2-P_S)P_S) - (1-P_S)^2(1-p-q)^{L-1}} \quad (65)$$

$$E\{T\} = \frac{1}{2p + q - 1} \times \frac{p^{L-1}((2p+q)^2 - P_S(3p+q-1) - p - q) - (1-P_S)(1-p-q)^{L-1}(L(2p+q-1) + p)}{p^{L-1}(P_S(p+q-1) - q(2p+q) + p+q) - (1-P_S)p(1-p-q)^{L-1}} \quad (66)$$

$$\tau = (1 - P_S) \frac{(1 - P_S)p(1 - p - q)^{L-1} + p^{L-1}(P_S(1 - p - q) + q(2p + q) - p - q)}{p^{L-1}((2 - P_S)P_S - 2p(1 - q) - (2 - q)q) + (1 - P_S)^2(1 - p - q)^{L-1}}. \quad (67)$$

**Case 2:** If link selection variable  $d_i$  is given by either (60) or (62), the probability of the buffer having  $k$  packets in its queue,  $\Pr\{Q = kR_0\}$ , is given by

$$\Pr\{Q = kR_0\} = \begin{cases} \frac{p^L(2p+q-1)}{p^L(2p+q-P_S)-(1-P_S)(1-p-q)^L}, & k = 0 \\ \frac{(1-P_S)(2p+q-1)p^{L-k}(1-p-q)^{k-1}}{p^L(2p+q-P_S)-(1-P_S)(1-p-q)^L}, & k = 1 \dots L \end{cases} \quad (68)$$

where, if link selection variable  $d_i$  is given by (60),  $p$  and  $q$  are given by (64), while if link selection variable  $d_i$  is given by (62),  $p$  and  $q$  are given by

$$p = 1 - P_R \quad \text{and} \quad q = P_S P_R + (1 - P_S)P_R P_C. \quad (69)$$



Furthermore, the average queue length,  $E\{Q\}$ , the average delay,  $E\{T\}$ , and throughput,  $\tau$ , are given by

$$E\{Q\} = R_0 \frac{1 - P_S}{2p + q - 1} \frac{p^{L+1} - (1 - p - q)^L (L(2p + q - 1) + p)}{p^L (2p + q - P_S) - (1 - P_S)(1 - p - q)^L} \quad (70)$$

$$E\{T\} = \frac{1}{2p + q - 1} \frac{1}{p} \frac{p^{L+1} - (1 - p - q)^L (L(2p + q - 1) + p)}{p^L - (1 - p - q)^L} \quad (71)$$

$$\tau = R_0(1 - P_S)p \frac{p^L - (1 - p - q)^L}{p^L (2p + q - P_S) - (1 - P_S)(1 - p - q)^L}. \quad (72)$$

*Proof:* Please refer to Appendix E. ■

Due to their complexity, the equations in Theorem 3 do not provide much insight into the performance of the considered system. To overcome this problem, we consider the case  $L \gg 1$ , which leads to significant simplifications and design insight. This is addressed in the following lemma.

**Lemma 6:** For the system considered in Theorem 3, assume that  $L \rightarrow \infty$ . In this case, for a system with link selection variable  $d_i$  given by (58), (60), or (62) to be able to achieve a fixed delay,  $E\{T\}$ , that does not grow with  $L$  as  $L \rightarrow \infty$ , the condition  $2p + q - 1 > 0$  must hold. If  $2p + q - 1 > 0$  holds, the following simplifications can be made for each of the considered link selection variables:

**Case 1:** If the link selection variable  $d_i$  is given by (58), the probability of the buffer being empty, the average delay,  $E\{T\}$ , and throughput,  $\tau$ , simplify to

$$\Pr\{Q = 0\} = P_S \frac{2P_C(1 - P_R)(1 - P_S) + (2 - P_R)P_S - 1}{2P_C(1 - P_S)(1 - P_S P_R) + P_S^2(1 - P_R)} \quad (73)$$

$$E\{T\} = \frac{1}{2P_C(1 - P_R)(1 - P_S) - P_R P_S + 2P_S - 1} + \frac{2P_C(1 - P_S)}{P_S^2(P_C(2P_R - 1) - P_R + 1) - 2P_C P_R P_S + P_C} \quad (74)$$

$$\tau = R_0(1 - P_S) \frac{P_S^2(P_C(2P_R - 1) - P_R + 1) - 2P_C P_R P_S + P_C}{2P_C(1 - P_S)(1 - P_S P_R) + (1 - P_R)P_S^2}. \quad (75)$$

**Case 2:** If the link selection variable  $d_i$  is given by (60), the probability of the buffer being empty, the average delay,  $E\{T\}$ , and throughput,  $\tau$ , simplify to

$$\Pr\{Q = 0\} = \frac{2P_C(1 - P_R)(1 - P_S) + P_S(2 - P_R) - 1}{(1 - P_R)(P_S + 2P_C(1 - P_S))} \quad (76)$$

$$E\{T\} = \frac{1}{2P_C(1 - P_R)(1 - P_S) - P_R P_S + 2P_S - 1} \quad (77)$$

$$\tau = R_0(1 - P_S) \frac{P_C(1 - P_S) + P_S}{2P_C(1 - P_S) + P_S}. \quad (78)$$

**Case 3:** If the link selection variable  $d_i$  is given by (62), the probability of the buffer being empty, the average delay,  $E\{T\}$ , and the throughput,  $\tau$ , simplify to

$$\Pr\{Q = 0\} = \frac{1 - P_R(2 - P_S - P_C(1 - P_S))}{2 - P_S - P_R(2 - P_S - P_C(1 - P_S))} \quad (79)$$

$$E\{T\} = \frac{1}{1 - P_R(2 - P_S - P_C(1 - P_S))} \quad (80)$$

$$\tau = R_0 \frac{1 + P_S P_R - P_R - P_S}{2 - P_S - P_R(2 - P_S - P_C(1 - P_S))}. \quad (81)$$

For each of the considered cases, the probability  $P_C$  can be used to adjust the desired average delay  $E\{T\}$  in (74), (77), and (80).

*Proof:* Please refer to Appendix F. ■

As already mentioned in Proposition 1, it is not possible to achieve any desired average delay with the proposed buffer-aided link selection protocols. The limits of the achievable average delay for each of the proposed link selection variables  $d_i$  in Proposition 1 are provided in the following lemma.

*Lemma 7:* Depending on the adopted link selection variable  $d_i$  the following cases can be distinguished for the average delay:

**Case 1:** If link selection variable  $d_i$  is given by (58), then if  $P_R < 1/(2 - P_S)$  and  $P_S < 1/(2 - P_R)$ , the system can achieve any average delay  $E\{T\} \geq T_{\min,1}$ , where  $T_{\min,1}$  is given by

$$T_{\min,1} = \frac{1}{1 - P_R(2 - P_S)} + \frac{2(1 - P_S)}{1 - P_S P_R(2 - P_S)}. \quad (82)$$

On the other hand, if  $P_R < 1/(2 - P_S)$  and  $P_S > 1/(2 - P_R)$ , the system can achieve any average delay in the interval  $T_{\min,1} \leq E\{T\} \leq T_{\max,1}$ , where  $T_{\max,1}$  is given by

$$T_{\max,1} = \frac{1}{P_S(2 - P_R) - 1}. \quad (83)$$

**Case 2:** If the link selection variable  $d_i$  is given by (60), then if  $P_R < 1/(2 - P_S)$  and  $P_S < 1/(2 - P_R)$ , the system can achieve any average delay  $E\{T\} \geq T_{\min,2}$ , where  $T_{\min,2}$  is given by

$$T_{\min,2} = \frac{1}{1 - P_R(2 - P_S)}. \quad (84)$$

However, if  $P_R < 1/(2 - P_S)$  and  $P_S > 1/(2 - P_R)$ , the system can achieve any average delay  $T_{\min,2} \leq E\{T\} \leq T_{\max,2}$ , where  $T_{\max,2} = T_{\max,1}$ .

**Case 3:** If the link selection variable  $d_i$  is given by (62), then if  $P_R > 1/(2 - P_S)$ , the system can achieve any average delay  $E\{T\} \geq T_{\min,3}$ , where  $T_{\min,3}$  is given by

$$T_{\min,3} = \frac{1}{1 - P_R}. \quad (85)$$

On the other hand, if  $P_R < 1/(2 - P_S)$ , the system can achieve any average delay  $T_{\min,3} \leq E\{T\} \leq T_{\max,3}$ , where  $T_{\max,3} = T_{\min,2}$ .

*Proof:* Please refer to Appendix G. ■

In the following, we investigate the outage probability of the proposed buffer-aided relaying protocol for delay constrained fixed rate transmission.

#### D. Outage Probability

The following theorem specifies the outage probability.

*Theorem 4:* For the considered buffer-aided relaying protocol in Proposition 1, if the required delay can be satisfied by using the link selection variable  $d_i$  in either (58) or (60), the outage probability is given by

$$\begin{aligned} F_{\text{out}} = & P_S \Pr\{Q = 0\} + P_S P_R (1 - \Pr\{Q = 0\} - \Pr\{Q = LR_0\}) \\ & + ((1 - P_S)P_R + (1 - P_S P_R)(1 - P_C)) \Pr\{Q = LR_0\}, \end{aligned} \quad (86)$$

where if  $d_i$  is given by (58),  $\Pr\{Q = 0\}$  and  $\Pr\{Q = LR_0\}$  are given by (63) with  $p$  and  $q$  given by (64). On the other hand, if  $d_i$  is given by (60),  $\Pr\{Q = 0\}$  and  $\Pr\{Q = LR_0\}$  are given by (68) with  $p$  and  $q$  given by (64).

If the required delay is satisfied by using the link selection variable  $d_i$  given by (62), then the outage probability is given by

$$\begin{aligned} F_{\text{out}} = & P_S \Pr\{Q = 0\} + P_S P_R (1 - \Pr\{Q = 0\} - \Pr\{Q = LR_0\}) \\ & + (1 - P_S)P_R(1 - P_C) \Pr\{Q = LR_0\}, \end{aligned} \quad (87)$$

where  $\Pr\{Q = 0\}$  and  $\Pr\{Q = LR_0\}$  are given by (60) with  $p$  and  $q$  given by (69).

*Proof:* Please refer to Appendix H. ■

The expressions for  $F_{\text{out}}$  in Theorem 4 are valid for general  $L$ . However, significant simplifications are possible if  $L \gg 1$ . This is addressed in the following lemma.

*Lemma 8:* When  $L \rightarrow \infty$ , the outages probability given by (86) and (87) simplifies to

$$F_{\text{out}} = P_S \Pr\{Q = 0\} + P_S P_R (1 - \Pr\{Q = 0\}) \quad (88)$$

where  $\Pr\{Q = 0\}$  is given by (73), (76), and (79) if  $d_i$  is given by (58), (60), and (62), respectively.

*Proof:* Eq. (88) is obtained by letting  $\Pr\{Q = LR_0\} \rightarrow 0$  when  $L \rightarrow \infty$  in (86) and (87). ■

The expression for the outage probability in (88) can be further simplified in the high SNR regime, which provides insight into the achievable diversity gain. This is summarized in the following theorem.

*Theorem 5:* In the high SNR regime, when  $\gamma_S = \gamma_R = \gamma \rightarrow \infty$ , depending on the required delay that the system has to satisfy, two cases can be distinguished:

**Case 1:** If  $1 < E\{T\} \leq 3$ , the outage probability asymptotically converges to

$$F_{\text{out}} \rightarrow \frac{P_S}{E\{T\} + 1}, \quad \text{as } \gamma \rightarrow \infty. \quad (89)$$

**Case 2:** If  $E\{T\} > 3$ , the outage probability asymptotically converges to

$$F_{\text{out}} \rightarrow \frac{P_S^2}{E\{T\} - 1} + P_S P_R, \quad \text{as } \gamma \rightarrow \infty. \quad (90)$$

Therefore, assuming Rayleigh fading, the considered system achieves a diversity gain of two if and only if  $E\{T\} > 3$ .

*Proof:* Please refer to Appendix I. ■

According to Theorem 5, for Rayleigh fading, a diversity gain of two can be also achieved for delay constrained transmission, which underlines the appeal of buffer-aided relaying with adaptive link selection compared to conventional relaying, which only achieves a diversity gain of one even in case of infinite delay (Conventional Relaying 1).

## V. MIXED RATE TRANSMISSION

In this section, we investigate buffer-aided relaying protocols with adaptive link selection for mixed rate transmission. In particular, we assume that the source does not have CSIT and transmits with fixed rate  $S_0$  but the relay has full CSIT and transmits with the maximum possible rate,  $R(i) = \log_2(1 + r(i))$ , that does not cause an outage in the  $\mathcal{R}\text{-}\mathcal{D}$  channel. For this scenario, we consider first delay unconstrained transmission and propose link adaptive buffer-aided relaying protocols with and without power allocation. Subsequently, we investigate the impact of delay constraints.

Before we proceed, we note that for mixed rate transmission the throughput can be expressed as

$$\tau = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N d_i \min\{\log_2(1 + r(i)), Q(i - 1)\}, \quad (91)$$

where we used (4) and (9). For the derivation of the maximum throughput of buffer-aided relaying with adaptive link selection the following theorem is useful.

*Theorem 6:* The link selection policy that maximizes the throughput of the considered buffer-aided relaying system for mixed rate transmission can be found in the set of link selection policies that satisfy

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N (1 - d_i) O_S(i) S_0 = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N d_i \log_2(1 + r(i)) . \quad (92)$$

Furthermore, for link selection policies within this set, the throughput is given by the right (or left) hand side of (92).

*Proof:* A proof of this theorem can be obtained by replacing  $O_R(i)R_0$  by  $\log_2(1 + r(i))$  in the proof of Theorem 1 given in Appendix A. ■

Hence, similar to fixed rate transmission, for the set of policies considered in Theorem 6, for  $N \rightarrow \infty$ , the buffer at the relay is practically always fully backlogged. Thus, the  $\min(\cdot)$  function in (91) can be omitted and the throughput is given by the right hand side of (92).

#### A. Optimal Link Selection Policy Without Power Allocation

Since the relay has the instantaneous CSI of both links, it can also optimize its transmit power. However, to get more insight, we first consider the case where the relay transmits with fixed power. We note that power allocation is not always desirable as it requires highly linear power amplifiers and thus, increases the implementation complexity of the relay.

According to Theorem 6, the optimal link selection policy maximizing the throughput can be found in the set of policies that satisfy (92). Therefore, the optimal policy can be obtained from the following optimization problem

$$\begin{aligned} \text{Maximize : } & \frac{1}{N} \sum_{i=1}^N d_i \log_2(1 + r(i)) \\ \text{Subject to : } & \text{C1 : } \frac{1}{N} \sum_{i=1}^N (1 - d_i) O_S(i) S_0 = \frac{1}{N} \sum_{i=1}^N d_i \log_2(1 + r(i)) \\ & \text{C2 : } \frac{1}{N} d_i (1 - d_i) = 0 \quad \forall i \end{aligned} \quad (93)$$

where  $N \rightarrow \infty$ , constraint C1 ensures that the search for the optimal policy is conducted only among the policies that satisfy (92), and C2 ensures that  $d_i \in \{0, 1\}$ . The solution of (93) leads to the following theorem.

*Theorem 7:* Let the pdfs of  $s(i)$  and  $r(i)$  be denoted by  $f_s(s)$  and  $f_r(r)$ , respectively. Then, for the considered buffer-aided relaying system in which the source transmits with a fixed rate  $S_0$  and fixed power  $\mathcal{P}_S$ , and the relay transmits with an adaptive rate  $R(i) = \log_2(1 + r(i))$  and fixed power  $\mathcal{P}_R$ , two

cases have to be distinguished for the optimal link selection variable  $d_i$ , which maximizes the throughput:

**Case 1:** If

$$P_S \leq \frac{S_0}{S_0 + \int_0^\infty \log_2(1+r) f_r(r) dr} \quad (94)$$

holds, then

$$d_i = \begin{cases} 1 & \text{if } O_S(i) = 0 \\ 1 & \text{if } O_S(i) = 1 \text{ AND } r(i) \geq 2^{\rho S_0} - 1 \\ 0 & \text{if } O_S(i) = 1 \text{ AND } r(i) < 2^{\rho S_0} - 1, \end{cases} \quad (95)$$

where  $\rho$  is a constant which can be found as the solution of

$$S_0(1 - P_S) \int_0^{2^{\rho S_0} - 1} f_r(r) dr = P_S \int_0^\infty \log_2(1+r) f_r(r) dr + (1 - P_S) \int_{2^{\rho S_0} - 1}^\infty \log_2(1+r) f_r(r) dr. \quad (96)$$

The maximum throughput is given by the right (or left) hand side of (96).

**Case 2:** If (94) does not hold, then

$$d_i = \begin{cases} 0 & \text{if } O_S(i) = 1 \\ 1 & \text{if } O_S(i) = 0. \end{cases} \quad (97)$$

In this case, the maximum throughput is given by

$$\tau = S_0(1 - P_S). \quad (98)$$

*Proof:* Please refer to Appendix J. ■

We note that with mixed rate transmission the  $\mathcal{S}\text{-}\mathcal{R}$  link is used only if it is not in outage, cf. (95), (97). On the other hand, the  $\mathcal{R}\text{-}\mathcal{D}$  is never in outage since the transmission rate is adjusted to the channel conditions. Furthermore, buffer-aided relaying with adaptive link selection has a larger throughput than Conventional Relaying 1, and also achieves a multiplexing gain of one.

To get more insight, we specialize the results derived thus far in this section to Rayleigh fading links.

*Lemma 9:* For Rayleigh fading links, condition (94) simplifies to

$$P_S = 1 - \exp\left(-\frac{2^{S_0} - 1}{\Omega_S}\right) \leq \frac{S_0}{S_0 + e^{1/\Omega_R} E_1(1/\Omega_R) / \ln(2)}. \quad (99)$$

Furthermore, (96) simplifies to

$$\begin{aligned} S_0 \exp\left(-\frac{2^{S_0} - 1}{\Omega_S}\right) \left[1 - \exp\left(-\frac{2^{\rho S_0} - 1}{\Omega_R}\right)\right] &= \frac{e^{1/\Omega_R}}{\ln(2)} \left[ \left(1 - \exp\left(-\frac{2^{S_0} - 1}{\Omega_S}\right)\right) E_1\left(\frac{1}{\Omega_R}\right) \right. \\ &\quad \left. + \exp\left(-\frac{2^{S_0} - 1}{\Omega_S}\right) E_1\left(\frac{2^{\rho S_0}}{\Omega_R}\right) \right] + \exp\left(-\frac{2^{\rho S_0} - 1}{\Omega_R}\right) \exp\left(-\frac{2^{S_0} - 1}{\Omega_S}\right) \rho S_0, \end{aligned} \quad (100)$$

and the maximum throughput is given by the right (or left) hand side of (100). If (99) does not hold, the throughput can be obtained by simplifying (98) to

$$\tau = S_0 \exp \left( -\frac{2^{S_0} - 1}{\Omega_S} \right). \quad (101)$$

*Proof:* Equations (99)-(101) are obtained by inserting the pdfs of  $s(i)$  and  $r(i)$  into (94), (96), and (98), respectively. ■

### B. Optimal Link Selection Policy with Power Allocation

As mentioned before, since for mixed rate transmission the relay is assumed to have the CSI of both links, power allocation can be applied to further improve performance. In other words, the relay can adjust its transmit power  $\mathcal{P}_R(i)$  to the channel conditions while the source still transmits with fixed power  $\mathcal{P}_S(i) = \mathcal{P}_S, \forall i$ . In the following, for convenience, we will use the transmit SNRs without fading,  $\gamma_S$  and  $\gamma_R(i)$ , which may be viewed as normalized powers, as variables instead of the actual powers  $\mathcal{P}_S = \gamma_S \sigma_{n_R}^2$  and  $\mathcal{P}_R(i) = \gamma_R(i) \sigma_{n_D}^2$ .

For the power allocation case, Theorem 6 is still applicable but it is convenient to rewrite the throughput as

$$\tau = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N d_i \log_2(1 + \gamma_R(i) h_R(i)). \quad (102)$$

We note that (92) also applies to the case of power allocation. Furthermore, in order to meet the average power constraint  $\Gamma$ , the instantaneous (normalized) power  $\gamma_R(i)$  and the fixed (normalized) power  $\gamma_S$  have to satisfy the following condition:

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N (1 - d_i) O_S(i) \gamma_S + \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N d_i \gamma_R(i) \leq \Gamma. \quad (103)$$

Thus, the optimal link selection policy for mixed rate transmission is the solution of the following optimization problem:

$$\begin{aligned} \text{Maximize : } & \frac{1}{N} \sum_{i=1}^N d_i \log_2(1 + \gamma_R(i) h_R(i)) \\ \text{Subject to : } & \text{C1 : } \frac{1}{N} \sum_{i=1}^N (1 - d_i) O_S(i) S_0 = \frac{1}{N} \sum_{i=1}^N d_i \log_2(1 + \gamma_R(i) h_R(i)) \\ & \text{C2 : } \frac{1}{N} d_i (1 - d_i) = 0, \quad \forall i \\ & \text{C3 : } \frac{1}{N} \sum_{i=1}^N (1 - d_i) O_S(i) \gamma_S + \frac{1}{N} \sum_{i=1}^N d_i \gamma_R(i) \leq \Gamma \end{aligned} \quad (104)$$



where  $N \rightarrow \infty$ , constraints C1 and C3 ensure that the search for the optimal policy is conducted only among those policies that jointly satisfy (92) and the source-relay power constraint (103), respectively, and C2 ensures that  $d_i \in \{0, 1\}$ . The solution of (104) is provided in the following theorem.

*Theorem 8:* Let the pdfs of  $h_S(i)$  and  $h_R(i)$  be denoted by  $f_{h_S}(h_S)$  and  $f_{h_R}(h_R)$ , respectively. Then, for the considered buffer-aided relaying system where the source transmits with a fixed rate  $S_0$  and fixed power  $\gamma_S$  and the relay transmits with adaptive rate  $R(i) = \log_2(1 + r(i)) = \log_2(1 + \gamma_R(i)h_R(i))$  and adaptive power  $\gamma_R(i)$ , two cases have to be considered for the optimal link selection variable  $d_i$  which maximizes the throughput:

**Case 1:** If

$$P_S \leq \frac{S_0}{S_0 + \int_{\lambda_t}^{\infty} \log_2(h_R/\lambda_t) f_{h_R}(h_R) dh_R}, \quad (105)$$

where  $\lambda_t$  is found as the solution to

$$P_S \int_{\lambda_t}^{\infty} \left( \frac{1}{\lambda_t} - \frac{1}{h_R} \right) f_{h_R}(h_R) dh_R + \gamma_S(1 - P_S) = \Gamma, \quad (106)$$

then the optimal power  $\gamma_R(i)$  and link selection variable  $d_i$  which maximize the throughput are given by

$$\gamma_R(i) = \max \left\{ 0, \frac{1}{\lambda} - \frac{1}{h_R(i)} \right\}, \quad (107)$$

and

$$d_i = \begin{cases} 1 & \text{if } O_S(i) = 0 \text{ AND } h_R(i) \geq \lambda \\ 1 & \text{if } O_S(i) = 1 \text{ AND } h_R(i) \geq \lambda \text{ AND } \ln \left( \frac{h_R(i)}{\lambda} \right) + \frac{\lambda}{h_R(i)} \geq \rho S_0 - \lambda \gamma_S + 1 \\ 0 & \text{if } O_S(i) = 1 \text{ AND } h_R(i) < \lambda \\ 0 & \text{if } O_S(i) = 1 \text{ AND } h_R(i) \geq \lambda \text{ AND } \ln \left( \frac{h_R(i)}{\lambda} \right) + \frac{\lambda}{h_R(i)} < \rho S_0 - \lambda \gamma_S + 1 \\ \varepsilon & \text{if } O_S(i) = 0 \text{ AND } h_R(i) < \lambda \end{cases} \quad (108)$$

where  $\varepsilon$  is either 0 or 1 and has not impact on the throughput. Constants  $\rho$  and  $\lambda$  are chosen such that constraints C1 and C3 in (104) are satisfied with equality. These two constants can be found as the solution to the following system of equations

$$S_0(1 - P_S) \int_0^G f_{h_R}(h_R) dh_R = P_S \int_{\lambda}^{\infty} \log_2 \left( \frac{h_R}{\lambda} \right) f_{h_R}(h_R) dh_R + (1 - P_S) \int_G^{\infty} \log_2 \left( \frac{h_R}{\lambda} \right) f_{h_R}(h_R) dh_R \quad (109)$$

$$\begin{aligned}
& P_S \int_{\lambda}^{\infty} \left( \frac{1}{\lambda} - \frac{1}{h_R} \right) f_{h_R}(h_R) dh_R + (1 - P_S) \int_G^{\infty} \left( \frac{1}{\lambda} - \frac{1}{h_R} \right) f_{h_R}(h_R) dh_R \\
& + \gamma_S (1 - P_S) \int_0^G f_{h_R}(h_R) dh_R = \Gamma
\end{aligned} \tag{110}$$

where the integral limit  $G$  is given by

$$G = -\frac{\lambda}{W\{-e^{\lambda\gamma_S - \rho S_0 - 1}\}}. \tag{111}$$

Here,  $W\{\cdot\}$  denotes the Lambert  $W$ -function defined in [26], which is available as built-in function in software packages such as Mathematica. The maximized throughput is given by the right (or left) hand side of (109).

**Case 2:** If (105) does not hold, the optimal power  $\gamma_R(i)$  and link selection variable  $d_i$  are given by

$$\gamma_R(i) = \max \left\{ 0, \frac{1}{\lambda} - \frac{1}{h_R(i)} \right\}, \text{ if } O_S(i) = 0; \tag{112}$$

$$d_i = \begin{cases} 0 & \text{if } O_S(i) = 1 \\ 1 & \text{if } O_S(i) = 0 \end{cases} \tag{113}$$

where  $\lambda = \lambda_t$  is the solution to (106). The maximum throughput is given by

$$\tau = S_0(1 - P_S). \tag{114}$$

*Proof:* Please refer to Appendix K. ■

*Remark 9:* Note that when conditions (94) and (105) do not hold, the throughput with and without power allocation is identical, cf. (98) and (114). If conditions (94) and (105) do not hold, this means that the SNR in the  $\mathcal{S}\text{-}\mathcal{R}$  channel is low, whereas the SNR in the  $\mathcal{R}\text{-}\mathcal{D}$  channel is high. In this case, power allocation is not beneficial since the  $\mathcal{S}\text{-}\mathcal{R}$  channel is the bottleneck link, which cannot be improved by power allocation at the relay. Furthermore, the throughput in (98) and (114) is identical to the throughput of a point-to-point communication between the source and the relay since the number of time slots required to transmit the information from relay to destination becomes negligible. Therefore, in this case, as far as the achievable throughput is concerned, the three-point half-duplex relay channel is transformed into a one hop channel between source and relay.

In the following lemma, we concentrate on Rayleigh fading for illustration purpose.

*Lemma 10:* For Rayleigh fading channels,  $P_S$  is given by

$$P_S = 1 - \exp \left( -\frac{2^{S_0} - 1}{\gamma_S \bar{\Omega}_S} \right).$$

Furthermore, condition (105) simplifies to

$$P_S \leq \frac{S_0}{S_0 + E_1(\lambda_t/\bar{\Omega}_R)/\ln(2)}, \quad (115)$$

where  $\lambda_t$  is found as the solution to

$$P_S \left[ \frac{e^{-\lambda_t/\bar{\Omega}_R}}{\lambda_t} - \frac{1}{\bar{\Omega}_R} E_1 \left( \frac{\lambda_t}{\bar{\Omega}_R} \right) \right] = \Gamma. \quad (116)$$

For the case where (115) holds, (109) and (110) simplify to

$$S_0(1 - P_S)(1 - e^{-G/\bar{\Omega}_R}) = \frac{1}{\ln(2)} \left[ P_S E_1 \left( \frac{\lambda}{\bar{\Omega}_R} \right) + (1 - P_S) \left( E_1 \left( \frac{G}{\bar{\Omega}_R} \right) + \ln \left( \frac{G}{\lambda} \right) e^{-G/\bar{\Omega}_R} \right) \right] \quad (117)$$

and

$$P_S \left[ \frac{e^{-\lambda/\bar{\Omega}_R}}{\lambda} - \frac{1}{\bar{\Omega}_R} E_1 \left( \frac{\lambda}{\bar{\Omega}_R} \right) \right] + (1 - P_S) \left[ \frac{e^{-G/\bar{\Omega}_R}}{\lambda} - \frac{1}{\bar{\Omega}_R} E_1 \left( \frac{G}{\bar{\Omega}_R} \right) \right] + \gamma_S(1 - P_S)(1 - e^{-G/\bar{\Omega}_R}) = \Gamma \quad (118)$$

respectively, where integral limit  $G$  is given by (111). The maximum throughput is given by the right (or left) hand side of (117).

For the case, where (115) does not hold, the throughput is given by  $\tau = S_0(1 - P_S)$ .

*Proof:* Equations (115), (116), (117), and (118) are obtained by inserting the pdfs of  $h_S(i)$  and  $h_R(i)$  into (105), (106), (109), and (110), respectively. ■

*Remark 10:* Conditions (94) and (105) depend only on the long term fading statistics and not on the instantaneous fading states. Therefore, for fixed  $\bar{\Omega}_S$  and  $\bar{\Omega}_R$ , the optimal policy for condition (94) is given by either (95) or (97), but not by both. Similarly, the optimal policy for condition (105) is given by either (108) or (113), but not by both.

### C. Mixed Rate Transmission with Delay Constraints

Now, we turn our attention to mixed rate transmission with delay constraints. For the delay unconstrained case, Theorem 6 was very useful to arrive at the optimal protocol since it removed the complexity of having to deal with the queue states. However, for the delay constrained case, the queue states determine the throughput and the average delay. Moreover, for mixed rate transmission, the queue states can only be modeled by a Markov chain with continuous state space, which makes the analysis complicated. Therefore, we resort to a suboptimal adaptive link selection protocol in the following.

*Proposition 2:* Let the buffer size be limited to  $Q_{\max}$  bits. For this case, we propose the following link selection protocol for mixed rate transmission with delay constraints:

- 1) If  $O_S(i) = 0$ , set  $d_i = 1$ .
- 2) Otherwise, if  $\log_2(1 + r(i)) \leq Q(i-1) \leq Q_{\max} - S_0$ , select  $d_i$  as proposed in Theorem 7 for the case of transmission without delay constraint.
- 3) Otherwise, if  $Q(i-1) > Q_{\max} - S_0$ , set  $d_i = 1$ .
- 4) Otherwise, if  $Q(i-1) < \log_2(1 + r(i))$ , set  $d_i = 0$ .

If the  $\mathcal{S}\text{-}\mathcal{R}$  link is in outage, the relay transmits. Otherwise, if there is enough room in the buffer to accommodate the bits possibly sent from the source to the relay and there are enough bits in the buffer for the relay to transmit, the link selection protocol introduced in Theorem 7 is employed. On the other hand, if there exists the possibility of a buffer overflow, the relay transmits to reduce the amount of data in the buffer. If the number of bits in the buffer is too low, the source transmits. The value of  $Q_{\max}$  can be used to adjust the average delay while maintaining a low throughput loss compared to the throughput without delay constraint.

Although conceptually simple, as pointed out before, a theoretical analysis of the throughput of the proposed queue size limiting protocol is difficult because of the continuous state space of the associated Markov chain. Thus, we will resort to simulations to evaluate its performance in Section VI.

#### D. Conventional Relaying With Delay Constraints

To have a benchmark for delay constrained buffer-aided relaying with adaptive link selection, we propose a corresponding conventional relaying protocol, which may be viewed as a delay constrained version of Conventional Relaying 1.

*Proposition 3:* The source transmits to the relay in  $k$  consecutive time slots followed by the relay transmitting to the destination in the following  $n$  time slots. Then, this pattern is repeated, i.e., the source transmits again in  $k$  consecutive time slots, and so on. The values of  $k$  and  $n$  can be chosen to satisfy any delay and throughput requirements.

For this protocol, the queue is non-absorbing if

$$k(1 - P_S)S_0 < nE\{\log_2(1 + r(i))\} \quad (119)$$

Assuming (119) holds, the average arrival rate is equal to the throughput and hence the throughput is given by

$$\tau = \frac{k}{k+n}(1 - P_S)S_0, \quad (120)$$

Using a numerical example, we will show in Section VI (cf. Fig. 7) that the protocol with adaptive link selection in Proposition 2 achieves a higher throughput than the conventional protocol in Proposition 3. However, the conventional protocol is more amenable to analysis and it is interesting to investigate the corresponding throughput and multiplexing gain for a given average delay in the high SNR regime,  $\gamma_S = \gamma_R = \gamma \rightarrow \infty$ . This is done in the following theorem.

*Theorem 9:* For a given average delay constraint,  $E\{T\}$ , the throughput  $\tau$  and multiplexing rate  $r$  of mixed rate transmission with conventional relaying are given by

$$\tau \rightarrow S_0 \left( 1 - \frac{1}{2E\{T\}} \right), \quad \text{as } \gamma \rightarrow \infty. \quad (121)$$

$$r \rightarrow 1 - \frac{1}{2E\{T\}}, \quad \text{as } \gamma \rightarrow \infty. \quad (122)$$

*Proof:* Please refer to Appendix L. ■

*Remark 11:* Theorem 9 reveals that, as expected from the discussion of the case without delay constraints, delay constrained mixed rate transmission approaches a multiplexing gain of one as the allowed average delay increases.

## VI. NUMERICAL AND SIMULATION RESULTS

In this section, we evaluate the performance of the proposed fixed rate and mixed rate transmission schemes. We also confirm some of our analytical results with computer simulations. We note that our analytical results are valid for  $N \rightarrow \infty$ . For the simulations,  $N$  has to be finite, of course, and transient effects caused by the filling and emptying of the buffer at the beginning and the end of transmission may have a non-negligible impact on the outage probability in this case. In order to avoid these transient effects in our simulations, we consider only bits transmitted during steady state operation, i.e., we neglect the bits transmitted at the beginning and the end of transmission.

### A. Fixed Rate Transmission

For fixed rate transmission, we evaluate the proposed link selection protocols for transmission with and without delay constraints. Throughout this section we assume that source and relay transmit with identical rates, i.e.,  $S_0 = R_0$ .

1) *Transmission Without Delay Constraints:* In Fig. 2, we show the ratio of the throughputs achieved with the proposed buffer-aided relaying protocol with adaptive link selection and Conventional Relaying 1 as a function of the transmit SNR  $\gamma_S = \gamma_R = \gamma$  for  $\bar{\Omega}_R = 1$ ,  $S_0 = R_0 = 2$  bits/slot, and different values of  $\bar{\Omega}_S$ . The throughput of buffer-aided relaying,  $\tau$ , was computed based on (36), (40), and (44) in Theorem 2, while the throughput of Conventional Relaying 1,  $\tau_{\text{conv},1}^{\text{fixed}}$ , was obtained based on (12). Furthermore, we also show simulation results where the throughput of the buffer-aided relaying protocol was obtained via Monte Carlo simulation. From Fig. 2 we observe that theory and simulation are in excellent agreement. Furthermore, Fig. 2 shows that except for  $\bar{\Omega}_S = \bar{\Omega}_R$  the proposed link adaptive relaying scheme achieves its largest gain for medium SNRs. For very high SNRs, both links are never in outage and thus, Conventional Relaying 1 with optimized  $\xi$  and link adaptive relaying achieve the same performance. On the other hand, for very low SNR, there are very few transmission opportunities on both links as the links are in outage most of the time. The proposed link adaptive protocol can exploit all of these opportunities. In contrast, for  $\bar{\Omega}_S = \bar{\Omega}_R$ , Conventional Relaying 1 choses  $\xi = 0.5$  and will miss half of the transmission opportunities by selecting the link that is in outage instead of the link that is not in outage because of the pre-determined schedule for link selection. On the other hand, if  $\bar{\Omega}_S$  and  $\bar{\Omega}_R$  differ significantly, Conventional Relaying 1 selects  $\xi$  close to 0 or 1 (depending on which link is stronger) and the loss compared to the link adaptive scheme becomes negligible.

In Fig. 3, we show the outage probability,  $F_{\text{out}}$ , for the proposed buffer-aided relaying protocol with adaptive link selection and Conventional Relaying 1. The same channel and system parameters as for Fig. 2 were adopted for Fig. 3 as well. For buffer-aided relaying with adaptive link selection,  $F_{\text{out}}$  was obtained from (51) and confirmed by Monte Carlo simulations. For conventional relaying,  $F_{\text{out}}$  was obtained from (13). As expected from Lemma 4, buffer-aided relaying achieves a diversity gain of two, whereas conventional relaying achieves only a diversity gain of one, which underlines the superiority of buffer-aided relaying with adaptive link selection.

2) *Transmission With Delay Constraints:* In Fig. 4, we show the throughput of buffer-aided relaying with adaptive link selection as function of the transmit SNR  $\gamma_S = \gamma_R = \gamma$  for fixed rate transmission with different constraints on the average delay  $E\{T\}$ . The theoretical curves for buffer-aided relaying were obtained from the throughput and the average delay given in Lemma 6. For comparison, we also show the throughput of buffer-aided relaying with adaptive link selection and without delay constraint (cf. Theorem 2), and the throughput of Conventional Relaying 2 given by (14). These two schemes introduce an infinite delay, i.e.,  $E\{T\} = \infty$ , and a delay of one time slot, respectively. In the low SNR

regime, the proposed buffer-aided relaying scheme with adaptive link selection cannot satisfy all delay requirements as expected from Lemma 7. Hence, for finite delays, the throughput curves in Fig. 4 do not extend to low SNRs. Nevertheless, as the affordable delay increases, the throughput for delay constrained transmission approaches the throughput for delay unconstrained transmission for sufficiently high SNR. Furthermore, the performance gain compared to Conventional Relaying 2 is substantial even for the comparatively small average delays  $E\{T\}$  considered in Fig. 4.

In Fig. 5, we show the outage probability,  $F_{\text{out}}$ , for the same schemes and parameters that were considered in Fig. 4. For buffer-aided relaying with adaptive link selection, the theoretical results shown in Fig. 5 were obtained from (86) and (87). These theoretical results are confirmed by the Monte Carlo simulation results also shown in Fig. 5. Furthermore, the curves for transmission without delay constraint (i.e.,  $E\{T\} = \infty$ ) were computed from (51), and for Conventional Relaying 2, we used (15). Fig. 5 shows that even for an average delay as small as  $E\{T\} = 1.1$  slots, buffer-aided relaying with adaptive link selection outperforms Conventional Relaying 2. Furthermore, as expected from Theorem 5, buffer-aided relaying with adaptive link selection achieves a diversity gain of two when the average delay is larger than three time slots (e.g.  $E\{T\} = 3.1$  time slots in Fig. 5). This leads to a large performance gain over conventional relaying which achieves only a diversity gain of one. Finally, note that even for  $E\{T\} = 3.1$  the coding gain loss is very small compared  $E\{T\} = \infty$ .

*Remark 12:* For the simulation results shown in Figs. 4 and 5, we adopted a relay with a buffer size of  $L = 60$  packets which leads to a negligible probability of dropped packets. For example, for  $\gamma = 45$  dB, the probability of a full buffer,  $\Pr\{Q = LR_0\}$ , is bounded by  $\Pr\{Q = LR_0\} < 10^{-60}$ . This also supports the claim in the proof of Theorem 5 that for large enough buffer sizes the probability of dropping a packet due to a buffer overflow becomes negligible.

### B. Mixed Rate Transmission

In this section, we investigate the achievable throughput for mixed rate transmission. For this purpose, we consider again the delay constrained and the delay unconstrained cases separately.

1) *Transmission Without Delay Constraints:* In Fig. 6, we compare the throughputs of buffer-aided relaying with adaptive link selection and Conventional Relaying 1. In both cases, we consider the cases with and without power allocation. The theoretical results shown in Fig. 6 for the four considered schemes were generated based on Theorem 7/Lemma 9, Theorem 8/Lemma 10, (21), (22), and (21), (23). The transmit SNRs of both links are identical, i.e.,  $\gamma_S = \gamma_R = \Gamma$ ,  $S_0 = 2$  bits/slot,  $\bar{\Omega}_S = 10$ , and  $\bar{\Omega}_R = 1$ . As



can be observed from Fig. 6, for both buffer-aided relaying with adaptive link selection and Conventional Relaying 1, power allocation is beneficial only for low to moderate SNRs. Both schemes can achieve a throughput of  $S_0$  bits/slot in the high SNR regime. However, adaptive link selection achieves a throughput gain compared to Conventional Relaying 1 in the entire considered SNR range.

2) *Transmission with Delay Constraints:* In Fig. 7, we compare the throughputs of various mixed rate and fixed rate transmission schemes for a maximum average delay of  $E\{T\} = 5$  time slots and  $S_0 = 2$  bits/slot. The transmit SNRs of both links are identical, i.e.,  $\gamma_S = \gamma_R = \Gamma$ ,  $\bar{\Omega}_S = \bar{\Omega}_R = 1$ . For mixed rate transmission, we simulated both the buffer-aided relaying protocol with adaptive link selection described in Proposition 2 and the conventional relaying described in Proposition 3. For fixed rate transmission, we chose  $R_0 = S_0 = 2$  bits/slot and included results for buffer-aided relaying with adaptive link selection obtained based on Lemma 6. Furthermore, for mixed rate transmission, we also show the maximum achievable throughput of buffer-aided relaying with adaptive link selection in the absence of delay constraints (as given by Theorem 7/Lemma 9) and the maximum throughput achievable for a delay constraint of  $E\{T\} = 5$  time slots and infinite transmit power (as given by (121)). Fig. 7 reveals that for mixed rate transmission the protocol with adaptive link selection proposed in Proposition 2 is superior to the conventional relaying scheme proposed in Proposition 3, and for high SNR, both protocols reach the upper bound for mixed rate transmission under a delay constraint given by (121). Furthermore, Fig. 7 also shows that mixed rate transmission is superior to fixed rate transmission since the former can exploit the additional flexibility afforded by having CSIT for the  $\mathcal{R}$ - $\mathcal{D}$  link. For example, for  $\Gamma = 30$  dB, mixed rate transmission with adaptive link selection achieves a throughput gain of 65 % compared to fixed rate transmission, and even conventional link selection still achieves a gain of 45%. Fig. 7 also shows that even in the presence of severe delay constraints mixed rate transmission can significantly reduce the throughput loss caused by half-duplexing compared to full-duplexing.

## VII. CONCLUSIONS

In this paper, we have considered a three-node DF relay system comprised of a source, a half-duplex relay with a buffer, and a destination, where the direct source-destination link is not available or not used. We have investigated both fixed rate transmission, where source and relay do not have CSIT and are forced to transmit with fixed rate, and mixed rate transmission, where the source does not have CSIT and transmits with fixed rate but the relay has full CSIT and transmits with variable rate. For both modes of transmission, we have derived the throughput-optimal buffer-aided relaying protocols with adaptive

link selection and the resulting throughputs and outage probabilities. Furthermore, we could show that buffer-aided relaying with adaptive link selection leads to substantial performance gains compared to conventional relaying with non-adaptive link selection. In particular, for fixed rate transmission, buffer-aided relaying with adaptive link selection achieves a diversity gain of two, whereas conventional relaying is limited to a diversity gain of one. For mixed rate transmission both buffer-aided relaying with adaptive link selection and a newly proposed conventional relaying scheme with non-adaptive link selection have been shown to overcome the half-duplex loss typical for wireless relaying protocols and to achieve a multiplexing gain of one. Since the proposed throughput-optimal protocols introduce an infinite delay, we have also proposed modified protocols for delay constrained transmission and have investigated the resulting throughput-delay trade-off. Surprisingly, the diversity gain of fixed rate transmission with buffer-aided relaying is also observed for delay constrained transmission as long as the average delay exceeds three time slots. Furthermore, for mixed rate transmission, for an average delay  $E\{T\}$ , a multiplexing gain of  $r = 1 - 1/(2E\{T\})$  is achieved even for conventional relaying.

## APPENDIX

### A. Proof of Theorem 1

We first note that, because of the law of the conservation of flow,  $A \geq \tau$  is always valid and equality holds if and only if the queue is non-absorbing.

We denote the set of indices with  $d_i = 1$  by  $\bar{I}$  and the set of indices with  $d_i = 0$  by  $I$ . Assume that we have a link selection protocol with arrival rate  $A$  and throughput  $\tau$  with  $A > \tau$ , i.e., the queue is absorbing. Then, for  $N \rightarrow \infty$ , we have

$$A = \frac{1}{N} \sum_{i \in I} (1 - d_i) O_S(i) S_0 > \tau = \frac{1}{N} \sum_{i \in \bar{I}} d_i O_R(i) \min\{R_0, Q(i-1)\}. \quad (123)$$

From (123) we observe that the considered protocol cannot be optimal as the throughput can be improved by moving some of the indices  $i$  in  $I$  to  $\bar{I}$  which leads to an increase of  $\tau$  at the expense of a decrease of  $A$ . As we continue moving indices from  $I$  to  $\bar{I}$  we reach a point where  $A = \tau$  holds. At this point, the queue becomes non-absorbing (but is at the boundary between a non-absorbing and an absorbing queue) and the throughput is maximized. If we continue moving indices from  $I$  to  $\bar{I}$ , in general,  $A$  will decrease and as a consequence of the law of conservation of flow,  $\tau$  will also decrease. We note that  $A$  does not decrease if we move only those indices from  $I$  to  $\bar{I}$  for which  $O_S(i) = 0$  holds. In this case,  $A$  will not change, and as a consequence of the law of conservation of flow, the value of  $\tau$  also remains

unchanged. However, the queue is moved from the edge of non-absorption if  $O_R(i) = 1$  holds for some of the indices moved from  $I$  to  $\bar{I}$ . However, as will be seen later, if the queue of the buffer operates at the edge of non-absorption, the throughput becomes independent of the state of the queue, which is desirable for analytical throughput maximization.

In the following, we will prove that when the queue is at the edge of non-absorption the following holds

$$\tau = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N d_i O_R(i) R_0 = A = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N (1 - d_i) O_S(i) S_0. \quad (124)$$

Let  $\epsilon$  denote a small subset of  $\bar{I}$  containing only indices  $i$  for which  $O_S(i) = 1$ , where  $|\epsilon|/N \rightarrow 0$  for  $N \rightarrow \infty$  and  $|\cdot|$  denotes the cardinality of a set. Throughout the remainder of this proof  $N \rightarrow \infty$  is assumed.

If the queue in the buffer of the relay is absorbing,  $A > \tau$  holds and on average the number of bits arriving at the queue exceed the number of bits leaving the queue. Thus,  $R_0 \leq Q(i-1)$  holds almost always and as a result the throughput can be written as

$$\tau = \frac{1}{N} \sum_{i \in \bar{I}} O_R(i) \min\{R_0, Q(i-1)\} = \frac{1}{N} \sum_{i \in \bar{I}} O_R(i) R_0. \quad (125)$$

Now, we assume that the queue is at the edge of non-absorption. That is  $A = \tau$  holds but moving the small fraction of indices in  $\epsilon$ , where  $|\epsilon|/N \rightarrow 0$ , from  $\bar{I}$  to  $I$  will make the queue an absorbing queue with  $A > \tau$ . For this case, we wish to determine whether or not

$$\frac{1}{N} \sum_{i \in \bar{I}} O_R(i) R_0 > \tau = \frac{1}{N} \sum_{i \in \bar{I}} O_R(i) \min\{R_0, Q(i-1)\} = A = \frac{1}{N} \sum_{i \in I} O_S(i) S_0 \quad (126)$$

holds. To test this, we move a small fraction  $\epsilon$ , where  $|\epsilon|/N \rightarrow 0$ , of indices from  $\bar{I}$  to  $I$ , thus making the queue an absorbing queue. As a result, (125) holds and (126) becomes

$$\frac{1}{N} \sum_{i \in \bar{I} \setminus \epsilon} O_R(i) R_0 = \tau = \frac{1}{N} \sum_{i \in \bar{I} \setminus \epsilon} O_R(i) \min\{R_0, Q(i-1)\} < A = \frac{1}{N} \sum_{i \in I \cup \epsilon} O_S(i) S_0. \quad (127)$$

From the above we conclude that if (125) holds, then based on (126) and (127), for  $|\epsilon|/N \rightarrow 0$ , we must have

$$\frac{1}{N} \sum_{i \in \bar{I}} O_R(i) R_0 > \frac{1}{N} \sum_{i \in I} O_S(i) S_0 \quad (128)$$

and

$$\frac{1}{N} \sum_{i \in \bar{I} \setminus \epsilon} O_R(i) R_0 < \frac{1}{N} \sum_{i \in I \cup \epsilon} O_S(i) S_0. \quad (129)$$

However, for (128) and (129) to jointly hold, we require that the particular considered move of indices from  $\bar{I}$  to  $I$  has caused a discontinuity in  $\frac{1}{N} \sum_{i \in \bar{I}} O_R(i) R_0$  or/and a discontinuity in  $\frac{1}{N} \sum_{i \in I} O_S(i) S_0$  as  $|\epsilon|/N \rightarrow 0$  is assumed. Since  $S_0$  and  $R_0$  are finite,  $\lim_{N \rightarrow \infty} \sum_{i \in \epsilon} S_0/N = \lim_{N \rightarrow \infty} S_0|\epsilon|/N = 0$  and  $\lim_{N \rightarrow \infty} \sum_{i \in \epsilon} R_0/N = \lim_{N \rightarrow \infty} R_0|\epsilon|/N = 0$ . Hence, such discontinuities are not possible. Therefore, at the edge of non-absorption the inequality in (126) cannot hold and we must have

$$\frac{1}{N} \sum_{i \in \bar{I}} O_R(i) R_0 = \tau = \frac{1}{N} \sum_{i \in \bar{I}} O_R(i) \min\{R_0, Q(i-1)\} = A = \frac{1}{N} \sum_{i \in I} O_S(i) S_0 \quad (130)$$

Eq. (130) can be written as (31). This concludes the proof.

### B. Proof of Theorem 2

The Lagrangian of Problem (32) is given by

$$\mathcal{L} = \frac{1}{N} \sum_{i=1}^N d_i O_R(i) R_0 - \mu \frac{1}{N} \sum_{i=1}^N [d_i O_R(i) R_0 - (1 - d_i) O_S(i) S_0] - \frac{1}{N} \sum_{i=1}^N \beta_i d_i (1 - d_i), \quad (131)$$

where  $\mu$  and  $\beta_i$  are the Lagrange multipliers. Differentiating  $\mathcal{L}$  with respect to  $d_i$  and setting the result to zero leads to

$$d_i = \frac{\beta_i + (-1 + \mu) O_R(i) R_0 + \mu O_S(i) S_0}{2\beta_i}. \quad (132)$$

For  $d_i(1 - d_i) = 0$  to hold, we need either  $d_i = 0$  or  $d_i = 1$ , which leads to two possible values for  $\beta_i$ :

$$d_i = 0 \Rightarrow \beta_{i,1} = (1 - \mu) O_R(i) R_0 - \mu O_S(i) S_0 \quad (133)$$

$$d_i = 1 \Rightarrow \beta_{i,2} = -\beta_{i,1} \quad (134)$$

For the maximum of  $\mathcal{L}$  in (131),  $\beta_i \leq 0, \forall i$ , has to hold. Hence, we have

$$d_i = \begin{cases} 1 & \text{if } (1 - \mu) O_R(i) R_0 \geq \mu O_S(i) S_0 \\ 0 & \text{if } (1 - \mu) O_R(i) R_0 \leq \mu O_S(i) S_0. \end{cases} \quad (135)$$

Furthermore,  $0 \leq \mu \leq 1$  has to hold since for  $\mu < 0$  and  $\mu > 1$  we have always  $d_i = 1$  and  $d_i = 0$ , respectively, irrespective of any non-negative values of  $O_S(i) S_0$  and  $O_R(i) R_0$ .

First, we consider the case  $0 < \mu < 1$ . The boundary values  $\mu = 0$  and  $\mu = 1$  will be investigated later. From (135), for  $0 < \mu < 1$ , we have four possibilities:

- 1) If  $O_R(i) = 1$  and  $O_S(i) = 0$ , then  $d_i = 1$ .
- 2) If  $O_R(i) = 0$  and  $O_S(i) = 1$ , then  $d_i = 0$ .

- 3) If  $O_R(i) = 0$  and  $O_S(i) = 0$ , then  $d_i$  can be chosen to be either  $d_i = 0$  or  $d_i = 1$  and the choice does not influence the throughput as both the source and the relay remain silent.
- 4) If  $O_R(i) = 1$  and  $O_S(i) = 1$  and  $\mu$  is chosen such that  $0 < \mu < R_0/(S_0 + R_0)$  then  $d_i = 1$  in all time slots with  $O_R(i) = 1$  and  $O_S(i) = 1$ , and as a result, condition C1 cannot be satisfied. Similarly, if  $\mu$  is chosen such that  $R_0/(S_0 + R_0) < \mu < 1$ , then  $d_i = 0$  in all time slots with  $O_R(i) = 1$  and  $O_S(i) = 1$ , and as a result condition C1 can also not be satisfied. Thus, we conclude that  $\mu$  must be set to  $\mu = R_0/(S_0 + R_0)$  since only in this case can  $d_i$  be chosen to be either  $d_i = 0$  or  $d_i = 1$ , which is necessary for satisfying condition C1. Since for  $O_R(i) = 1$  and  $O_S(i) = 1$  neither link is in outage,  $d_i$  can be chosen to be either zero or one, as long as condition C1 is satisfied. In order to satisfy C1, we propose to flip a coin and the outcome of the coin toss decides whether  $d_i = 1$  or  $d_i = 0$ . Let the coin have two outcomes  $\mathcal{C} \in \{0, 1\}$  with probabilities  $\Pr\{\mathcal{C} = 0\}$  and  $\Pr\{\mathcal{C} = 1\}$ . We set  $d_i = 0$  if  $\mathcal{C} = 0$  and  $d_i = 1$  if  $\mathcal{C} = 1$ . Thus, the probabilities  $\Pr\{\mathcal{C} = 0\}$  and  $\Pr\{\mathcal{C} = 1\}$  have to be chosen such that C1 is satisfied.

Choosing the link selection variable as in (34) and exploiting the independence of  $s(i)$  and  $r(i)$ , condition C1 results in

$$S_0 [(1 - P_S)P_R + (1 - P_S)(1 - P_R)\Pr\{\mathcal{C} = 0\}] = R_0 [(1 - P_R)P_S + (1 - P_S)(1 - P_R)\Pr\{\mathcal{C} = 1\}]. \quad (136)$$

From (136), we can obtain the probabilities  $\Pr\{\mathcal{C} = 0\}$  and  $\Pr\{\mathcal{C} = 1\}$ , which after some basic algebraic manipulations leads to (35). The throughput is given by the right (or left) hand side of (136), which leads to (36).

For (35) to be valid,  $\Pr\{\mathcal{C} = 0\}$  and  $\Pr\{\mathcal{C} = 1\}$  have to meet  $0 \leq \Pr\{\mathcal{C} = 0\} \leq 1$  and  $0 \leq \Pr\{\mathcal{C} = 1\} \leq 1$ , which leads to the conditions

$$S_0(1 - P_S) - (1 - P_R)P_S R_0 \geq 0 \quad (137)$$

$$R_0(1 - P_R) - (1 - P_S)P_R S_0 \geq 0 \quad (138)$$

Solving (137) and (138), we obtain that for the link selection variable  $d_i$  given in (34) to be valid, condition (33) has to be fulfilled.

Next, we consider the case where  $\mu = 0$ . Inserting  $\mu = 0$  in (135), we obtain three possible cases:

- 1) If  $O_R(i) = 1$ , then  $d_i = 1$ .
- 2) If  $O_R(i) = 0$  and  $O_S(i) = 0$ , then  $d_i$  can be chosen to be either  $d_i = 0$  or  $d_i = 1$  and the choice has no influence on the throughput.

- 3) If  $O_R(i) = 0$  and  $O_S(i) = 1$ , then  $d_i$  can be chosen to be either  $d_i = 0$  or  $d_i = 1$  as long as condition C1 is satisfied. Similar to before, in order to satisfy C1, we propose to flip a coin and the outcome of the coin flip determines whether  $d_i = 1$  or  $d_i = 0$ .

Choosing the link selection variable as in (38) and exploiting the independence of  $s(i)$  and  $r(i)$ , condition C1 can be rewritten as

$$S_0 P_R (1 - P_S) \Pr\{\mathcal{C} = 0\} = R_0 (1 - P_R). \quad (139)$$

After basic manipulations (139) simplifies to (39). The throughput is given by the right (or left) hand side of (139) and can be simplified to (40). Imposing again the conditions  $0 \leq \Pr\{\mathcal{C} = 0\} \leq 1$  and  $0 \leq \Pr\{\mathcal{C} = 1\} \leq 1$ , we find that for  $\mu = 0$ , (137) still has to hold but (138) can be violated, which is equivalent to the new condition

$$P_R > \frac{R_0}{R_0 + S_0(1 - P_S)}. \quad (140)$$

Finally, letting  $\mu = 1$  and following a similar path as for  $\mu = 0$  leads to (42)–(44) and condition (41). This concludes the proof.

### C. Proof of Lemma 2

We provide two different proofs for the outage probability,  $F_{\text{out}}$ , in (51). The first proof is more straightforward and based on (10). However, the second proof provides more insight into when outages occur.

*Proof 1:* In the absence of outages, the maximum achievable throughput, denoted by  $\tau_0$ , is given by (50). Thus, when (33) holds,  $F_{\text{out}}$  is obtained by inserting (36) and (50) into (10). Similarly, when (37) holds,  $F_{\text{out}}$  is obtained by inserting (40) and (50) into (10). Finally, when (41) holds,  $F_{\text{out}}$  is obtained by inserting (44) and (50) into (10). After basic simplifications, (51) is obtained. This concludes the proof.

*Proof 2:* The second proof exploits the fact that an outage occurs when both the source and the relay are silent, i.e., when none of the links is used. When (33) holds, from  $d_i$  given by (34), we observe that no transmission occurs only when both links are in outage. This happens with probability  $F_{\text{out}} = P_S P_R$ . In contrast, when (37) holds, from  $d_i$  given by (38), we observe that no node transmits when both links are in outage or when the  $\mathcal{S}$ - $\mathcal{R}$  link is not in outage, while the  $\mathcal{R}$ - $\mathcal{D}$  link is in outage and the coin flip chooses the relay for transmission. This event happens with probability  $F_{\text{out}} = P_S P_R + (1 - P_S) P_R P_C$ , which after inserting  $P_C$  given by (39) leads to (51). Finally, when (41) holds, from  $d_i$ , given by (42),

we see that no node transmits when both links are in outage or when the  $\mathcal{S}\text{-}\mathcal{R}$  link is in outage, while the  $\mathcal{R}\text{-}\mathcal{D}$  link is not in outage and the coin flip chooses the source for transmission. This happens with probability  $F_{\text{out}} = P_S P_R + P_S(1 - P_R)(1 - P_C)$ , which after introducing  $P_C$  given by (43) leads to (51).

#### D. Proof of Lemma 4

Computing the link outages in (7) and (8) for Rayleigh fading and exploiting (53), we obtain (54) by employing  $\Omega_S = \gamma \bar{\Omega}_S$  and  $\Omega_R = \gamma \bar{\Omega}_R$  in the resulting expression and using a Taylor series expansion for  $\gamma \rightarrow \infty$ . As can be seen from (54), the transmit SNR  $\gamma$  has an exponent of  $-2$ . Thus, the diversity order is two.

Moreover, for  $\bar{\Omega}_S = \bar{\Omega}_R = \bar{\Omega}$  and  $S_0 = R_0$ , the asymptotic expression for  $F_{\text{out}}$  in (54) simplifies to

$$F_{\text{out}} \rightarrow \frac{(2^{R_0} - 1)^2}{\bar{\Omega}^2 \gamma^2}, \quad \text{as } \gamma \rightarrow \infty. \quad (141)$$

Furthermore, for  $S_0 = R_0$ , the asymptotic throughput in (52) simplifies to  $\tau = R_0/2$ . Thus, letting  $\tau = r \log_2(1 + \gamma)$  we obtain  $R_0 = 2r \log_2(1 + \gamma)$ . Inserting  $R_0 = 2r \log_2(1 + \gamma)$  into (141), the diversity-multiplexing trade-off,  $DM(r)$ , is obtained as

$$\begin{aligned} DM(r) &= - \lim_{\gamma \rightarrow \infty} \frac{\log_2(F_{\text{out}})}{\log_2(\gamma)} = - \lim_{\gamma \rightarrow \infty} \frac{2 \log_2(2^{2r \log_2(1+\gamma)} - 1) - 2 \log_2(\bar{\Omega}) - 2 \log_2(\gamma)}{\log_2(\gamma)} \\ &= 2 - \lim_{\gamma \rightarrow \infty} \frac{2 \log_2((1 + \gamma)^{2r} - 1)}{\log_2(\gamma)} = 2 - 4r. \end{aligned} \quad (142)$$

This completes the proof.

#### E. Proof of Theorem 3

Let  $d_i$  be given by (58). Then, the following events are possible for the queue in the buffer:

- 1) If the buffer is empty, it stays empty with probability  $P_S$  and receives one packet with probability  $1 - P_S$ .
- 2) If the buffer contains one packet, it stays in the same state with probability  $P_S P_R$ , sends the packet with probability  $P_S(1 - P_R)$ , and receives a new packet with probability  $1 - P_S$ .
- 3) If the buffer contains more than one packet but less than  $L$  packets, it stays in the same state with probability  $P_S P_R$ , receives a new packet with probability  $(1 - P_S)P_R + (1 - P_S)(1 - P_R)(1 - P_C)$ , and sends one packet with probability  $(1 - P_R)P_S + (1 - P_S)(1 - P_R)P_C$ .
- 4) If the buffer contains  $L$  packets, it stays in the same state with probability  $P_S P_R + (1 - P_S)P_R + (1 - P_S)(1 - P_R)(1 - P_C)$ , and sends one packet with probability  $(1 - P_R)P_S + (1 - P_S)(1 - P_R)P_C$ .

The events for the queue of the buffer detailed above, form a Markov chain whose states are defined by the number of packets in the queue. This Markov chain is shown in Fig. 8, where the probabilities  $p$  and  $q$  are given by (64). Let  $\mathbf{M}$  denote the state transition matrix of the Markov chain and let  $m_{i,j}$  denote the element in the  $i$ -th row and  $j$ -th column of  $\mathbf{M}$ . Then,  $m_{i,j}$  is the probability that the buffer will transition from having  $i - 1$  packets in its queue in the previous time slot to having  $j - 1$  packets in its queue in the following time slot. The non-zero elements of matrix  $\mathbf{M}$  are given by

$$\begin{aligned} m_{1,1} &= P_S, \quad m_{1,2} = 1 - P_S, \quad m_{2,1} = P_S - q, \quad m_{2,3} = 1 - P_S, \quad m_{L+1,L+1} = 1 - p \\ m_{i,i+1} &= 1 - p - q, \quad m_{i+1,i} = p, \quad m_{i,i} = q, \quad \text{for } i = 1, \dots, L. \end{aligned} \quad (143)$$

Let  $\Pr\{\mathbf{Q}\} = [\Pr\{Q = 0\}, \Pr\{Q = R_0\}, \dots, \Pr\{Q = LR_0\}]$  denote the steady state probability vector of the considered Markov chain, where  $\Pr\{Q = kR_0\}$ ,  $k = 0, \dots, L$ , is the probability of having  $k$  packets in the buffer. The steady state probability vector is obtained by solving the following system of equations

$$\begin{cases} \Pr\{\mathbf{Q}\}\mathbf{M} &= \Pr\{\mathbf{Q}\} \\ \sum_{k=0}^L \Pr\{Q = kR_0\} &= 1 \end{cases}, \quad (144)$$

which leads to (63). Using (63) the average queue size  $E\{Q\}$  can be obtained from

$$E\{Q\} = R_0 \sum_{k=0}^L k \Pr\{Q = kR_0\}, \quad (145)$$

which leads to (65). Furthermore, the average arrival rate can be found as

$$\begin{aligned} A &= R_0 [(1 - P_S)(\Pr\{Q = 0\} + \Pr\{Q = R_0\}) \\ &\quad + (1 - p - q)(1 - \Pr\{Q = 0\} - \Pr\{Q = R_0\} - \Pr\{Q = LR_0\})]. \end{aligned} \quad (146)$$

Inserting the average arrival rate given by (146) and the average queue size given by (65) into (56) yields the average delay in (66).

For the case when  $d_i$  is given by either (60) or (62), the queue in the buffer of the relay can be modeled by the Markov chain shown in Fig. 9. If the link selection variable  $d_i$  is given by (60),  $p$  and  $q$  are given by (64), and if the link selection variable  $d_i$  is given by (62),  $p$  and  $q$  are given by (69). Following the same procedure as before, (68)-(72) can be obtained. This completes the proof.



### F. Proof of Lemma 6

Let us first assume that  $2p + q - 1 < 0$ , which is equivalent to  $p < 1 - p - q$ . Now, since  $L \rightarrow \infty$ ,  $p^L$  goes to zero faster than  $(1 - p - q)^L$ . Thus, by using  $p^L = 0$  as  $L \rightarrow \infty$  in (66) and (71), we obtain in both cases

$$E\{T\} = \frac{L}{p} - \frac{1}{1 - 2p - q}. \quad (147)$$

Thus, we conclude that if  $2p + q - 1 < 0$ ,  $E\{T\}$  grows with  $L$  and is unlimited as  $L \rightarrow \infty$ . Thus, if  $E\{T\}$  is to be limited as  $L \rightarrow \infty$ ,  $2p + q - 1 > 0$  has to hold.

If  $2p + q - 1 > 0$ , as  $L \rightarrow \infty$ ,  $(1 - p - q)^L$  goes to zero faster than  $p^L$ . Hence, (73)-(81) are obtained by letting  $(1 - p - q)^L = 0$ , as  $L \rightarrow \infty$ , in the relevant equations in Theorem 3 and inserting the corresponding  $p$  and  $q$  given by (64) and (69) into the resulting expressions. This concludes the proof.

### G. Proof of Lemma 7

The minimum and maximum possible delays that the considered buffer-aided relaying system can achieve are obtained for  $P_C = 1$  and  $P_C = 0$ , respectively. If  $d_i$  is given by (58), the delay is given by (74). By setting  $P_C = 1$  in (74) we obtain the minimum possible delay in (82). However, since (74) is valid only when  $2p + q - 1 > 0$ , (82) is valid only when  $P_R < 1/(2 - P_S)$ . This condition is obtained by inserting  $P_C = 1$  into the expressions for  $p$  and  $q$  given by (64) and exploiting  $2p + q - 1 > 0$ . On the other hand, in order to get the maximum delay given in (83), we set  $P_C = 0$  in (74). The derived maximum delay is valid only when  $P_S > 1/(2 - P_R)$ , which is obtained from  $2p + q - 1 > 0$  and inserting  $P_C = 0$  into the expressions for  $p$  and  $q$  given by (64).

A similar approach can be used to derive the delay limits  $T_{\min,2}$ ,  $T_{\max,2}$ ,  $T_{\min,3}$ , and  $T_{\max,3}$  valid for the cases when  $d_i$  is given by (60) and (62). This concludes the proof.

### H. Proof of Theorem 4

The outage probability,  $F_{\text{out}}$ , can be derived based on two different approaches. The first approach is straightforward and based on (10). However, the second approach provides more insight into how and when the outages occur and is based on counting the time slots in which no transmissions occur. In the following, we provide a proof based on the latter approach.

If  $d_i$  is given by (58) or (60), there are four different cases where no node transmits.

- 1) The buffer is empty and the  $\mathcal{S}$ - $\mathcal{R}$  link is in outage.

- 2) The buffer is not empty nor full and both the  $\mathcal{S}\text{-}\mathcal{R}$  and  $\mathcal{R}\text{-}\mathcal{D}$  links are in outage.
- 3) The buffer is full and the  $\mathcal{S}\text{-}\mathcal{R}$  link is not in outage while the  $\mathcal{R}\text{-}\mathcal{D}$  link is in outage. In this case, the source is selected for transmission but since the buffer is full, the packet is dropped.
- 4) The buffer is full, both the  $\mathcal{S}\text{-}\mathcal{R}$  and  $\mathcal{R}\text{-}\mathcal{D}$  links are not in outage, and the source is selected for transmission based on the coin flip. In this case, since the buffer is full, the packet is dropped.

Summing up the probabilities for each of the above four cases, we obtain (86).

If  $d_i$  is given by (62), an outage occurs in three cases: Case 1 and Case 2 as described above, and a new Case 3. In the new Case 3, the buffer is full, the  $\mathcal{S}\text{-}\mathcal{R}$  link is not in outage while the  $\mathcal{R}\text{-}\mathcal{D}$  link is in outage, and the source is selected for transmission based on the coin flip. Summing up the probabilities for each of the three cases, we obtain (87).

### I. Proof of Theorem 5

For delay constrained transmission with  $E\{T\} < L$ , the probability of dropped packets  $\Pr\{Q = LR_0\}$  can be made arbitrarily small by increasing the buffer size  $L$ . Thus, for large enough  $L$ , we can set  $\Pr\{Q = LR_0\} = 0$  in (86) and (87).

In the high SNR regime, when  $P_S \rightarrow 0$  and  $P_R \rightarrow 0$ ,  $P_R < 1/(2 - P_S)$  and  $P_S < 1/(2 - P_R)$  always hold. Using  $P_S \rightarrow 0$  and  $P_R \rightarrow 0$  in the delays specified in Proposition 1, we obtain the conditions  $E\{T\} > 3$  and  $1 < E\{T\} \leq 3$  if link selection variable  $d_i$  is given by (58) and (60), respectively.

We first consider the case  $E\{T\} > 3$ , where  $d_i$  is given by (58). Thus, the probability of the buffer being empty,  $\Pr\{Q = 0\}$ , is given by (73). Using  $P_S \rightarrow 0$  and  $P_R \rightarrow 0$  in (73), we obtain

$$\Pr\{Q = 0\} = P_S \left(1 - \frac{1}{2P_C}\right). \quad (148)$$

On the other hand, using  $P_S \rightarrow 0$  and  $P_R \rightarrow 0$  in the expression for  $E\{T\}$  in (74), we obtain

$$E\{T\} = \frac{1}{2P_C - 1} + 2. \quad (149)$$

Solving (149) for  $P_C$  yields

$$P_C = \frac{1}{2} \left(1 + \frac{1}{E\{T\} - 2}\right). \quad (150)$$

Inserting (150) into (148) we obtain

$$\Pr\{Q = 0\} = \frac{P_S}{E\{T\} - 1}. \quad (151)$$

Finally, inserting (151) into (86) and setting  $\Pr\{Q = LR_0\} = 0$ , we obtain (90).

Now, we consider the case  $1 < E\{T\} \leq 3$ , where  $d_i$  is given by (60). Here, the probability of the buffer being empty,  $\Pr\{Q = 0\}$ , is given by (76). For  $P_S \rightarrow 0$  and  $P_R \rightarrow 0$ , we obtain from (76)

$$\Pr\{Q = 0\} = 1 - \frac{1}{2\Pr\{\mathcal{C} = 1\}}. \quad (152)$$

Furthermore, for  $P_S \rightarrow 0$  and  $P_R \rightarrow 0$ , we obtain from (77) the asymptotic delay

$$E\{T\} = \frac{1}{2P_C - 1} \quad (153)$$

or equivalently

$$P_C = \frac{1}{2} \left( 1 + \frac{1}{E\{T\}} \right). \quad (154)$$

Inserting (154) into (152) we obtain

$$\Pr\{Q = 0\} = \frac{1}{E\{T\} + 1}. \quad (155)$$

Finally, inserting (155) into (87) and setting  $\Pr\{Q = LR_0\} = 0$ , we obtain (89). This concludes the proof.

#### J. Proof of Theorem 7

The Lagrangian of optimization problem (93) is given by

$$\begin{aligned} \mathcal{L} = & \frac{1}{N} \sum_{i=1}^N d_i \log_2 (1 + r(i)) - \mu \frac{1}{N} \sum_{i=1}^N \left[ d_i \log_2 (1 + r(i)) \right. \\ & \left. - (1 - d_i) O_S(i) S_0 \right] - \frac{1}{N} \sum_{i=1}^N \beta_i d_i (1 - d_i), \end{aligned} \quad (156)$$

where  $\mu$  and  $\beta_i$  are the Lagrange multipliers. By differentiating  $\mathcal{L}$  with respect to  $d_i$ , equating the result to zero, and solving the equation with respect to  $d_i$ , we obtain

$$d_i = \begin{cases} 1 & \text{if } (1 - \mu) \log_2 (1 + r(i)) \geq \mu O_S(i) S_0 \\ 0 & \text{if } (1 - \mu) \log_2 (1 + r(i)) \leq \mu O_S(i) S_0, \end{cases} \quad (157)$$

where we took into account that  $\beta_i < 0$ . Since for  $\mu < 0$  and  $\mu > 1$ , we have always  $d_i = 1$  and  $d_i = 0$ , respectively, irrespective of the (non-negative) values of  $\log_2 (1 + r(i))$  and  $O_S(i) S_0$ ,  $0 \leq \mu \leq 1$  has to hold.

Let us first consider the case  $0 < \mu < 1$  and investigate the boundary values  $\mu = 0$  and  $\mu = 1$  later. For  $0 < \mu < 1$ , (157) can be written in the form of (95) after setting  $\rho = \mu/(1 - \mu)$ , where  $\rho$  is chosen such that constraint C1 of problem (93) is met. Denoting the pdfs of  $s(i)$  and  $r(i)$  by  $f_s(s)$  and  $f_r(r)$

constraint C1 of problem (93) can be rewritten as in (96), which is valid for  $\rho$  in the range of  $\rho = [0, \infty)$ . Thus, by setting  $\rho = \infty$  in (96), we obtain the entire domain over which (95) is valid, which leads to condition (94).

Next, we consider the boundary values  $\mu = 0$  and  $\mu = 1$ .  $\mu = 0$  or equivalently  $\rho = 0$  is relevant only in the trivial case when the  $\mathcal{S}$ - $\mathcal{R}$  link is never in outage (i.e.  $P_S = 0$ ) and  $S_0 = \infty$ , where a trivial solution is given by  $d_1 = 0$  and  $d_i = 1$  for  $i = 2, \dots, N$  and  $N \rightarrow \infty$ .

The other boundary value,  $\mu = 1$ , is invoked only when by using  $d_i$  as defined in (95), constraint C1 cannot be satisfied even when  $\rho \rightarrow \infty$ , which is the case when condition (94) does not hold. Therefore, if (94) does not hold, we set  $\mu = 1$  in (157) and obtain the following cases:

- 1) If  $O_S(i) = 1$ , then  $d_i = 0$ .
- 2) If  $O_S(i) = 0$ , then  $d_i$  can be chosen arbitrarily to be either zero or one as long as constraint C1 holds.

However, the same throughput as obtained when  $O_S(i) = 0$  and  $d_i$  is chosen such that constraint C1 holds, can also be obtained by choosing always  $d_i = 1$  when  $O_S(i) = 0$  resulting in (97). The reason behind this is as follows: Assume there is a policy for which when  $O_S(i) = 0$ ,  $d_i$  is chosen such that constraint C1 holds. Now, we change  $d_i$  from 0 to 1 for  $O_S(i) = 0$ . However, this change does not affect the (average) amount of data entering the buffer. Thus, because of the law of conservation of flow, the average amount of data entering the buffer per time slot is identical to the average amount of data leaving the buffer per time slot (the throughput), and the throughput is not affected by the change.

#### K. Proof of Theorem 8

The Lagrangian of optimization problem (104) is given by

$$\begin{aligned} \mathcal{L} = & \frac{1}{N} \sum_{i=1}^N d_i \log_2(1 + \gamma_R(i)h_R(i)) - \mu \frac{1}{N} \sum_{i=1}^N \left[ d_i \log_2(1 + \gamma_R(i)h_R(i)) - (1 - d_i)O_S(i)S_0 \right] \\ & - \nu \frac{1}{N} \sum_{i=1}^N \left[ (1 - d_i)O_S(i)\gamma_S + d_i\gamma_R(i) \right] - \frac{1}{N} \sum_{i=1}^N \beta_i d_i (1 - d_i), \end{aligned} \quad (158)$$

where the Lagrange multipliers  $\mu$ ,  $\beta_i$ , and  $\nu$  are chosen such that C1, C2, and C3 are satisfied, respectively. We again consider only the interval  $0 \leq \mu \leq 1$  as for  $\mu < 0$  and  $\mu > 1$ , we have always  $d_i = 1$  and  $d_i = 0$ , respectively, irrespective of the (nonnegative) values of  $\log_2(1 + r(i))$  and  $O_S(i)S_0$ .

We concentrate first on the case  $0 < \mu < 1$  and consider the boundary values later. By differentiating  $\mathcal{L}$  with respect to  $\gamma_R(i)$  and  $d_i$ , and setting the results to zero, we obtain two equations. Solving the

resulting system of equations with respect to  $\gamma_R(i)$  and  $d_i$ , and taking into account that  $\beta_i < 0$ ,  $0 < \mu < 1$ , and  $\nu > 0$ , we obtain (107) and (108) after letting  $\rho = \ln(2)\mu/(1 - \mu)$  and  $\lambda = \ln(2)\nu/(1 - \mu)$ , which are chosen such that constraints C1 and C3 are met with equality. Given the pdfs  $f_{h_S}(h_S)$  and  $f_{h_R}(h_S)$ , conditions (92) and (103) can be directly written as (109) and (110), respectively. Setting  $\rho \rightarrow \infty$  in (109) and (110), we obtain condition (105) which is necessary for the validity of (97).

Similar to the fixed transmit power case, the boundary value  $\mu = 0$  is trivial. On the other hand, for  $\mu = 1$ , we obtain that  $d_i$  has to be set to  $d_i = 0$  when  $O_S(i) = 1$  and for  $O_S(i) = 0$ ,  $d_i$  can be chosen arbitrarily. Similar to the fixed power case, we set  $d_i = 1$  when  $O_S(i) = 1$  in order to minimize the delay. Thus, the optimal power and link selection variables are given by (112) and (113), respectively, and the throughput is given by (114).

#### L. Proof of Theorem 9

For  $\gamma_S = \gamma_R = \gamma \rightarrow \infty$ , the protocol in Proposition 3 is optimal in the sense that it maximizes the throughput while satisfying the average delay constraint. In particular, for high SNR in the  $\mathcal{S}\text{-}\mathcal{R}$  link, the probability that the link is in outage approaches zero and the relay receives  $S_0$  bits per source transmission. On the other hand, the number of bits transmitted by the relay in one time slot over the  $\mathcal{R}\text{-}\mathcal{D}$  link increases with the SNR. Thus, for sufficiently high SNR, the source transmits  $kS_0$  bit in  $k$  time slots and the relay needs just  $n = 1$  time slot to forward the information to the destination. Hence, every transmission period comprises  $k + n = k + 1$  time slots, where the queue length at the relay increases from  $S_0$  to  $kS_0$  in the first  $k$  time slots and is reduced to zero in the  $(k + 1)$ th time slot. Hence, the average queue length,  $E\{Q\}$ , can be written as

$$E\{Q\} \rightarrow \frac{1}{k+1}(1 + 2 + \dots + k + 0)S_0 = \frac{1}{k+1} \frac{k(k+1)}{2} S_0 = \frac{k}{2} S_0, \quad \text{as } \gamma \rightarrow \infty. \quad (159)$$

On the other hand, the arrival rate is identical to the throughput and given by (120), and for high SNR it converges to

$$A = \tau \rightarrow S_0 \frac{k}{k+1}, \quad \text{as } \gamma \rightarrow \infty. \quad (160)$$

Combining (56), (159), and (160) the average delay is found as

$$E\{T\} \rightarrow \frac{k+1}{2}, \quad \text{as } \gamma \rightarrow \infty. \quad (161)$$

Finally, combining (160) and (161) the throughput can be expressed as (121), and the multiplexing gain in (122) follows directly.

## REFERENCES

- [1] E. C. V. D. Meulen, "Three-Terminal Communication Channels," *Advances in Applied Probability*, vol. 3, pp. 120–154, 1971.
- [2] T. Cover and A. El Gamal, "Capacity Theorems for the Relay Channel," *IEEE Trans. Inform. Theory*, vol. 25, pp. 572–584, Sep. 1979.
- [3] A. Host-Madsen and J. Zhang, "Capacity Bounds and Power Allocation for Wireless Relay Channels," *IEEE Trans. Inform. Theory*, vol. 51, pp. 2020–2040, Jun. 2005.
- [4] A. Sendonaris, E. Erkip, and B. Aazhang, "User Cooperation Diversity — Part I. System Description," *IEEE Trans. Commun.*, vol. 51, pp. 1927–1938, Nov. 2003.
- [5] —, "User Cooperation Diversity — Part II. Implementation Aspects and Performance Analysis," *IEEE Trans. Commun.*, vol. 51, pp. 1939–1948, Nov. 2003.
- [6] J. Laneman, D. Tse, and G. Wornell, "Cooperative Diversity in Wireless Networks: Efficient Protocols and Outage Behavior," *IEEE Trans. Inform. Theory*, vol. 50, pp. 3062–3080, Dec. 2004.
- [7] J. Laneman and G. Wornell, "Distributed Space-Time Block Coded Protocols for Exploiting Cooperative Diversity in Wireless Networks," *IEEE Trans. Inform. Theory*, vol. 49, pp. 2415–2425, Oct. 2003.
- [8] A. Nosratinia, T. Hunter, and A. Hedayat, "Cooperative Communication in Wireless Networks," *IEEE Communications Magazine*, vol. 42, pp. 74–80, Oct. 2004.
- [9] T. Hunter and A. Nosratinia, "Diversity Through Coded Cooperation," *IEEE Trans. Wireless Commun.*, vol. 5, pp. 283–289, 2006.
- [10] K. Azarian, H. E. Gamal, and P. Schniter, "On the Achievable Diversity-Multiplexing Tradeoff in Half-Duplex Cooperative Channels," *IEEE Trans. Inform. Theory*, vol. 51, pp. 4152–4172, Dec. 2005.
- [11] R. Nabar, H. Bölcskei, and F. Kneubühler, "Fading Relay Channels: Performance Limits and Space-Time Signal Design," *IEEE J. Select. Areas Commun.*, vol. 22, pp. 1099–1109, Aug. 2004.
- [12] A. Bletsas, A. Khisti, D. Reed, and A. Lippman, "A Simple Cooperative Diversity Method Based on Network Path selection," *IEEE J. Select. Areas Commun.*, vol. 24, pp. 659–672, Mar. 2006.
- [13] F. Xue and S. Sandhu, "Cooperation in a Half-Duplex Gaussian Diamond Relay Channel," *IEEE Trans. Inform. Theory*, vol. 53, pp. 3806–3814, Oct. 2007.
- [14] G. Kramer, M. P. Gastpar, and Gupta, "Cooperative Strategies and Capacity Theorems for Relay Networks," *IEEE Trans. Inform. Theory*, vol. 51, pp. 3037–3063, Sep. 2005.
- [15] Y. Liang, V. Veeravalli, and H. Poor, "Resource Allocation for Wireless Fading Relay Channels: Max-Min Solution," *IEEE Trans. Inform. Theory*, vol. 53, pp. 3432–3453, Oct. 2007.
- [16] B. Xia, Y. Fan, J. Thompson, and H. V. Poor, "Buffering in a Three-Node Relay Network," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 4492–4496, Nov. 2008.
- [17] A. Ikhlef, D. Michalopoulos, and R. Schober, "Max-Max Relay Selection for Relays with Buffers," *IEEE Trans. Wireless Commun.*, vol. 11, pp. 1124 – 1135, Mar. 2012.
- [18] R. Wang, V. Lau, and K. Huang, "A New Scaling Law on Throughput and Delay Performance of Wireless Mobile Relay Networks Over Parallel Fading Channels," in *IEEE International Symposium on Information Theory*, Jul. 2009.
- [19] Q. Wang, P. Fan, M. McKay, and K. Letaief, "On the Position Selection of Relays in Diamond Relay Networks," *IEEE Trans. Commun.*, vol. 59, pp. 2515–2527, 2011.
- [20] L. Ding, M. Tao, F. Yang, and W. Zhang, "Joint Scheduling and Relay Selection in One- and Two-Way Relay Networks with Buffering," in *IEEE International Conference on Communications, ICC '09.*, Jun. 2009.

- [21] N. Zlatanov, R. Schober, and P. Popovski, "Buffer-Aided Relaying with Adaptive Link Selection," *Accepted for publication in IEEE J. Select. Areas Commun.*, 2012. [Online]. Available: <http://www.ece.ubc.ca/~zlatanov/pdf/JSAC2011.pdf>
- [22] —, "Throughput and Diversity Gain of Buffer-Aided Relaying," in *IEEE Global Telecommunications Conference (GLOBECOM 2011)*, Dec. 2011.
- [23] L. L. Yang, C. Dong, and L. Hanzo, "Multihop Diversity - A Precious Source of Fading Mitigation in Multihop Wireless Networks," in *IEEE Global Telecommunications Conference (GLOBECOM 2011)*, Dec. 2011.
- [24] L. Georgiadis, M. J. Neely, and L. Tassiulas, *Resource Allocation and Cross Layer Control in Wireless Networks*. Now Publishers Inc, 2006.
- [25] J. D. C. Little, "A Proof of the Queueing Formula:  $L = \lambda w$ ," *Operations Research*, vol. 9, no. 3, pp. 383–388, May - Jun. 1961.
- [26] R. Corless, G. Gonnet, D. Hare, D. Jeffrey, and D. Knuth, "On the LambertW function," *Advances in Computational Mathematics*, vol. 5, no. 1, pp. 329–359, 1996.

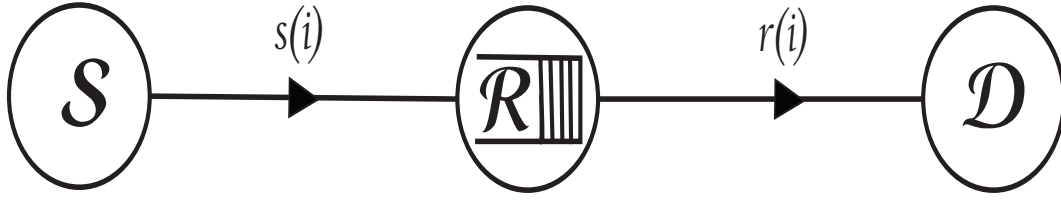


Fig. 1. System model for three node relay network employing a half-duplex decode-and-forward relay. The relay is equipped with a buffer to store the packets received from the source.

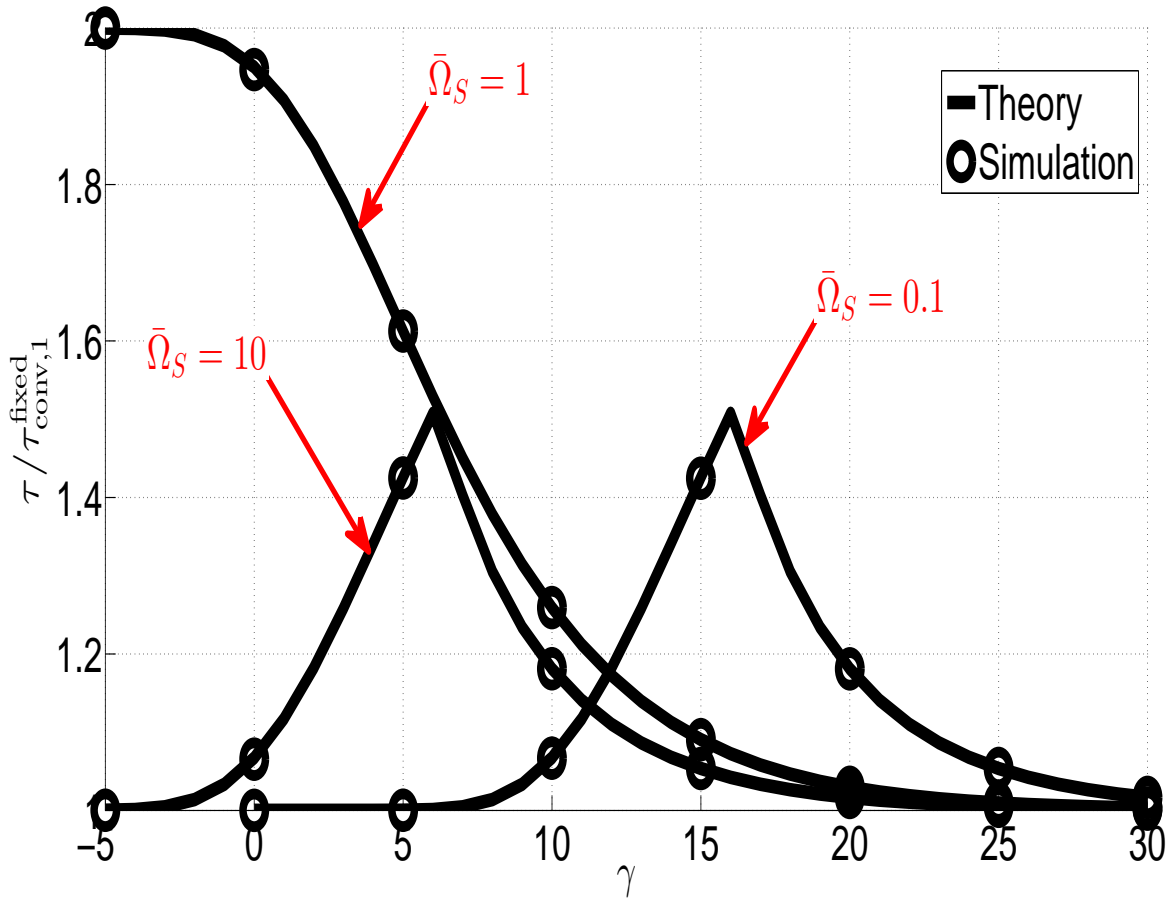


Fig. 2. Ratio of the throughputs of buffer-aided relaying and Conventional Relaying 1,  $\tau / \tau_{\text{conv},1}^{\text{fixed}}$ , vs.  $\gamma$ . Fixed rate transmission without delay constraints.  $\gamma_S = \gamma_R = \gamma$ ,  $S_0 = R_0 = 2$  bits/slot, and  $\bar{\Omega}_R = 1$ .



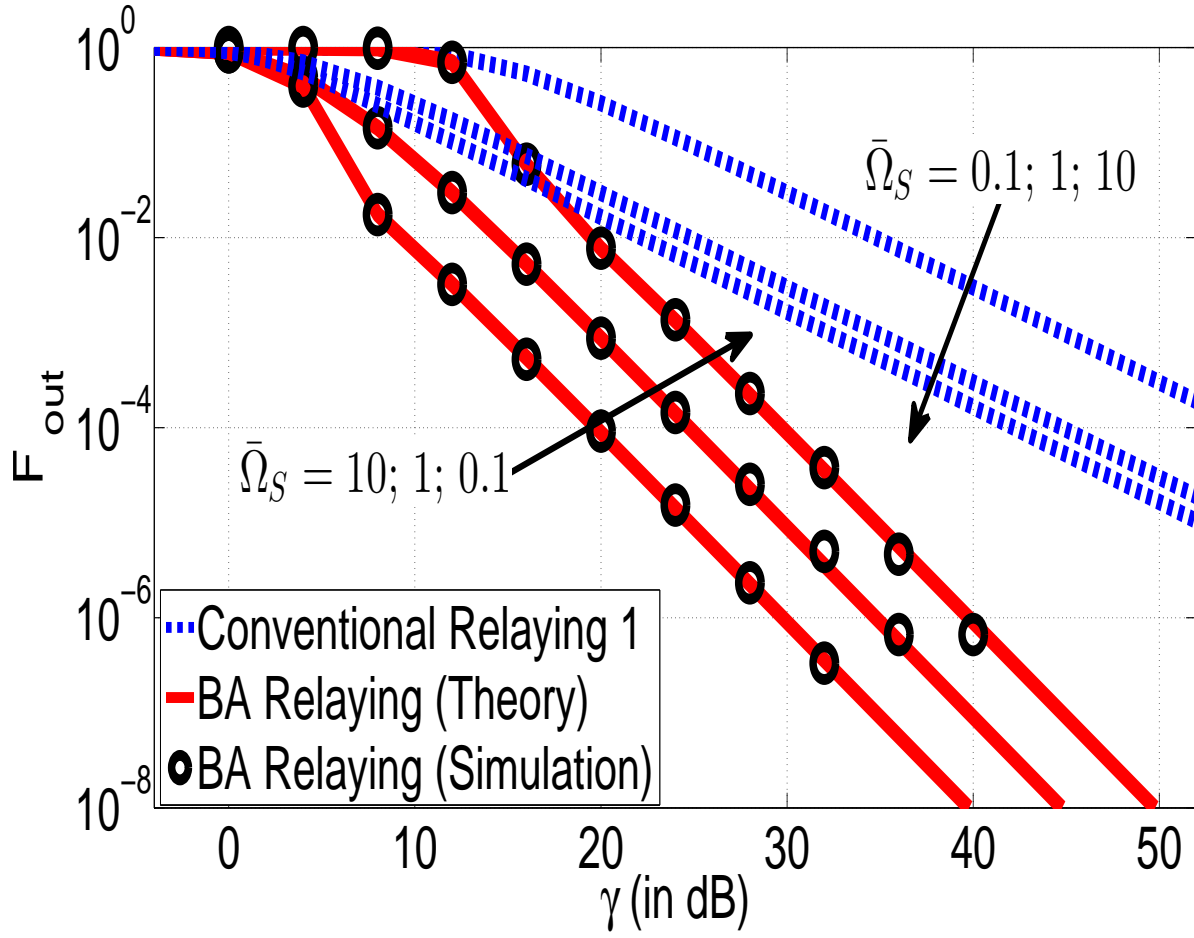


Fig. 3. Outage probability of buffer-aided (BA) relaying and Conventional Relaying 1 vs.  $\gamma$ . Fixed rate transmission without delay constraints.  $\gamma_S = \gamma_R = \gamma$ ,  $S_0 = R_0 = 2$  bits/slot, and  $\bar{\Omega}_R = 1$ .

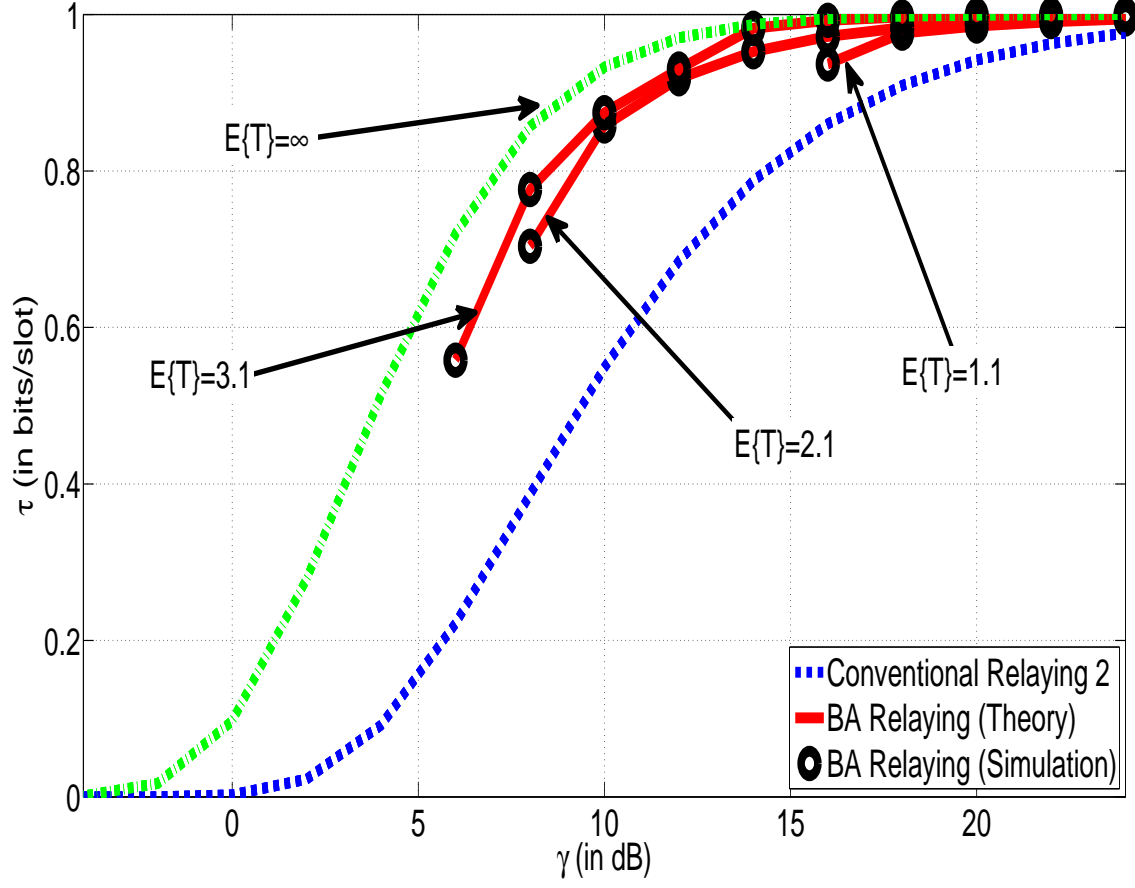


Fig. 4. Throughputs of buffer-aided (BA) relaying and Conventional Relaying 2 vs.  $\gamma$ . Fixed rate transmission with delay constraints.

$\gamma_S = \gamma_R = \gamma$ ,  $S_0 = R_0 = 2$  bits/slot,  $\bar{\Omega}_R = 1$ , and  $\bar{\Omega}_S = 1$ .

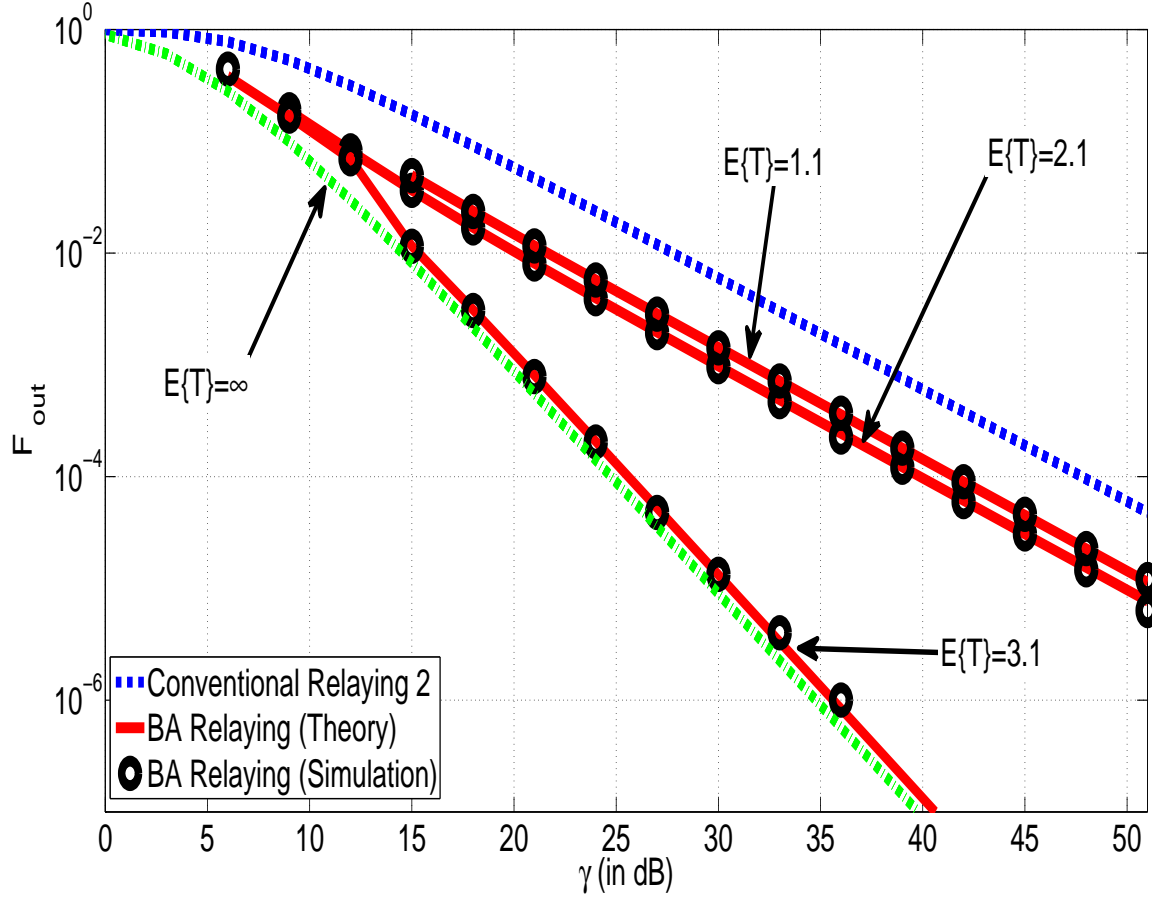


Fig. 5. Outage probability of buffer-aided (BA) relaying and Conventional Relaying 2 vs.  $\gamma$ . Fixed rate transmission with delay constraints.

$\gamma_S = \gamma_R = \gamma$ ,  $S_0 = R_0 = 2$  bits/slot,  $\bar{\Omega}_R = 1$ , and  $\bar{\Omega}_S = 1$ .

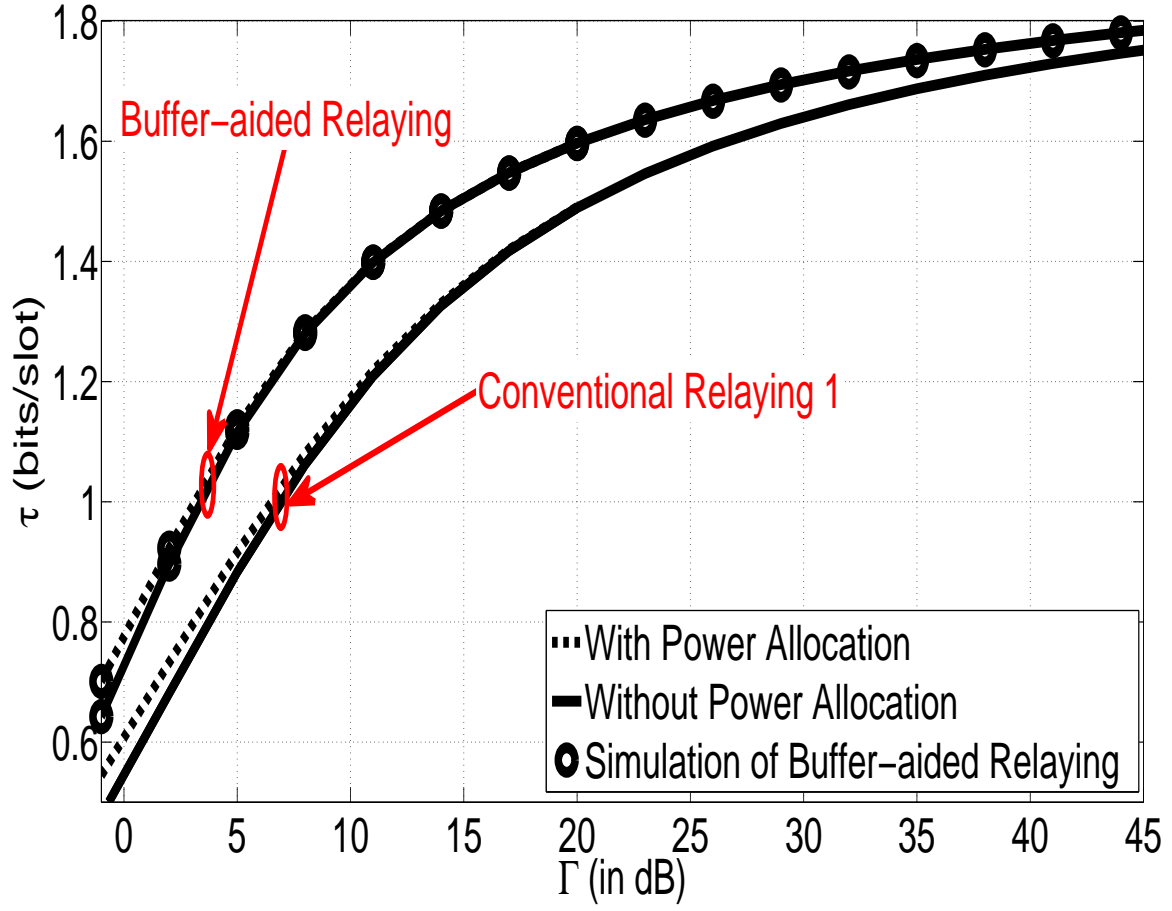


Fig. 6. Throughput of buffer-aided relaying with adaptive link selection and Conventional Relaying 1 vs.  $\Gamma$ . Mixed rate transmission without delay constraints.  $\bar{\Omega}_S = 10$ ,  $\bar{\Omega}_R = 1$ , and  $S_0 = 2$  bits/slot.

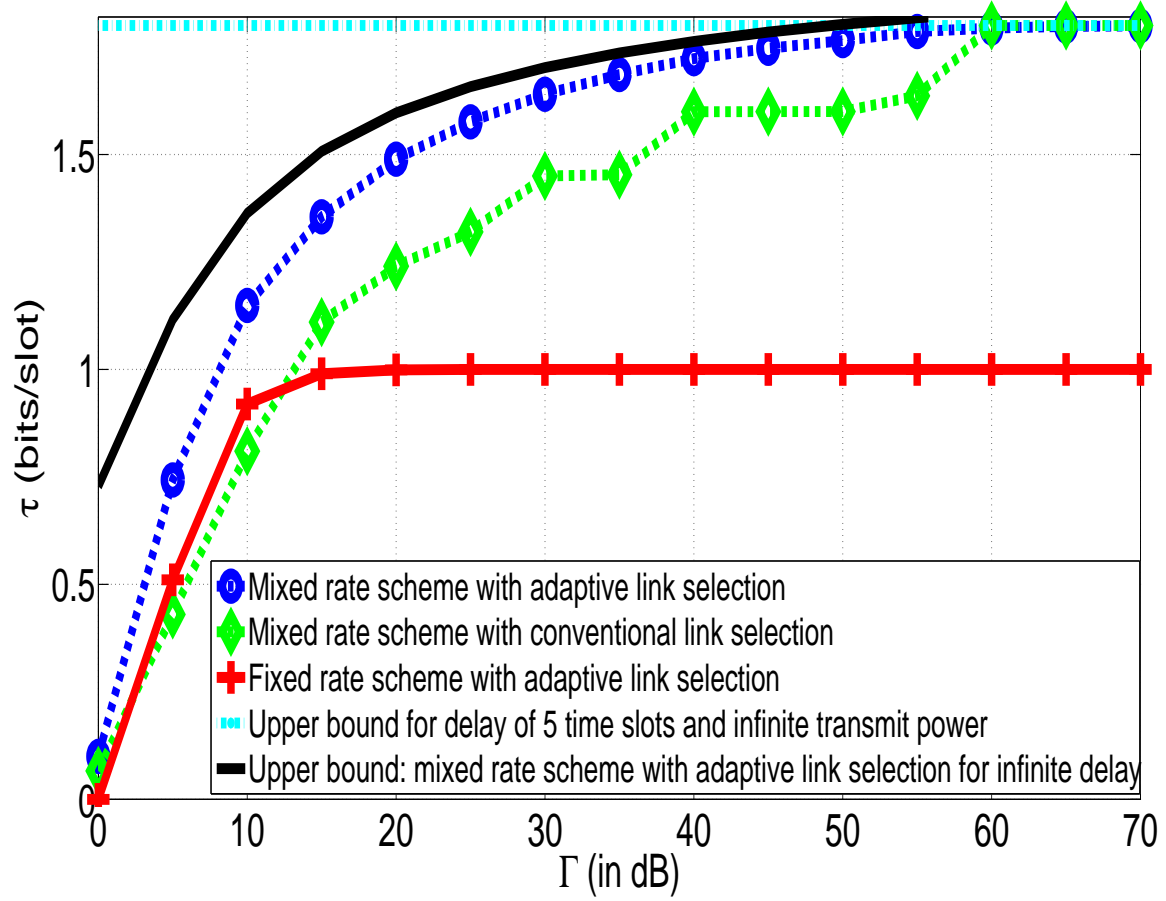


Fig. 7. Throughput of buffer-aided relaying with adaptive link selection and conventional relaying vs.  $\Gamma$ . Mixed rate and fixed rate transmission with delay constraint.  $E\{T\} = 5$  time slots,  $\gamma_S = \gamma_R = \Gamma$ ,  $S_0 = 2$  bits/slot, and  $\bar{\Omega}_S = \bar{\Omega}_R = 1$ .

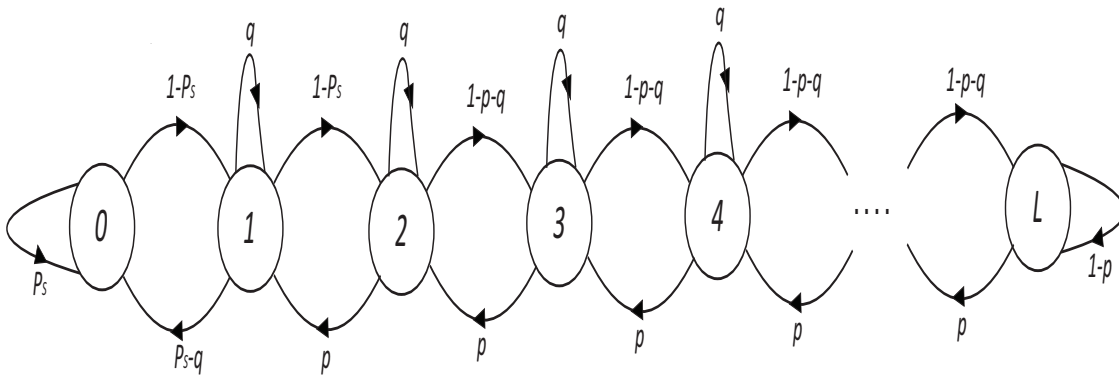


Fig. 8. Markov chain for the number of packets in the queue of the buffer if the link selection variable  $d_i$  is given by (58).

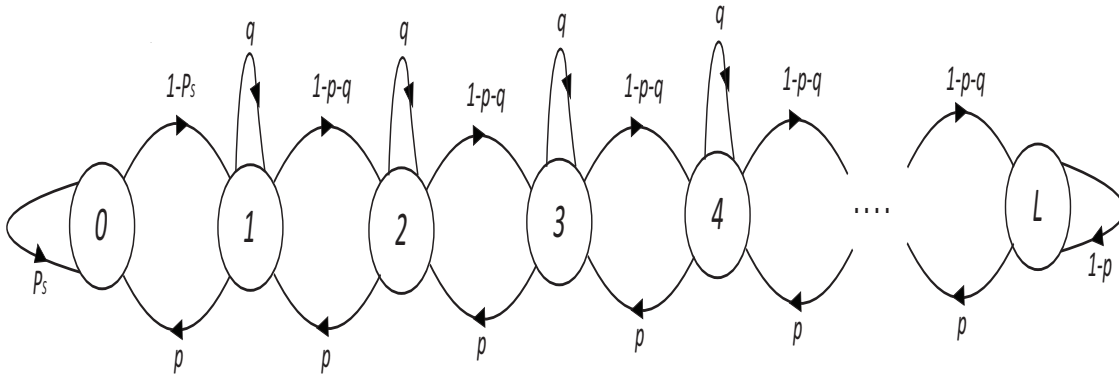


Fig. 9. Markov chain for the number of packets in the queue of the buffer if the link selection variable  $d_i$  is given by (60) or (62).